

Gencore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 09:56:37 ; Search time 30.56 Seconds

Perfect score: 1065 (without alignments)
Sequence: 1 MAKQPSDVSSCDEGRQLQ PRMVLRLRLRYTVRLVWRMH 198

Title: US-09-508-832-10
Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 74754 seqs, 110103796 residues

Total number of hits satisfying chosen parameters:

74754

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : A_Geneseq_032802:*

1: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1980.DAT:/*
2: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1981.DAT:/*
3: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1982.DAT:/*
4: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1983.DAT:/*
5: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1984.DAT:/*
6: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1985.DAT:/*
7: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1986.DAT:/*
8: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1987.DAT:/*
9: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1988.DAT:/*
10: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1989.DAT:/*
11: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1990.DAT:/*
12: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1991.DAT:/*
13: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1992.DAT:/*
14: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1993.DAT:/*
15: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1994.DAT:/*
16: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1995.DAT:/*
17: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1996.DAT:/*
18: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1997.DAT:/*
19: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1998.DAT:/*
20: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA1999.DAT:/*
21: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA2000.DAT:/*
22: /SIDS1/gcadata/hold-geneseq/geneseqp-emb/AA2001.DAT:/*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Location/Qualifiers
1	1065	100.0	198	20 AAW98158	Human Bcl-2 inter	FT Key Binding-site 42..131 "dynein light chain binding region"
2	908	85.3	196	20 AAW98166	Human Bcl-2 inter	FT Region 92..100 "note" "BH3"
3	696	65.4	138	20 AAW98164	Human Bcl-2 mutant	FT Region 121..135 "note" "hydrophobic region"
4	696	65.4	138	20 AAW98157	Human Bcl-2 inter	
5	691	64.9	138	20 AAW98165	Human Bcl-2 mutant	
6	691	64.9	138	20 AAW98166	Human Bcl-2 mutant	
7	691	64.9	138	20 AAW98168	Human Bcl-2 mutant	
8	690	64.8	138	20 AAW98167	Human Bcl-2 mutant	
9	589	55.3	140	20 AAW98165	Human Bcl-2 inter	PR 24..SEP..1997; 97AU-0009373.
10	584	54.8	140	20 AAW98160	Human Bcl-2 inter	PR 17..SEP..1997; 97AU-0009263.
11	584	54.8	140	20 AAW98161	Human Bcl-2 mutant	

12	584	54.8	140	20 AAW98163	Murine Bim-L mutant
13	583	54.7	140	20 AAW98162	Murine Bim-L mutant
14	582	54.6	140	20 AAW98159	Murine Bim-L mutant
15	497	46.7	92	22 ABG01163	Novel human diagno.
16	406	38.1	110	20 AAW98154	Murine Bcl-2 inter
17	144	13.5	27	21 AAB37019	BC12 polypeptide B
18	138	13.0	26	21 AAY97352	Mammalian Bim Bcl-3 consensus
19	130	13.0	26	21 AAB070378	Human polypeptide
20	123	11.5	23	22 AA009353	Human polypeptide
21	114	10.7	157	22 ARO06556	Human duncel-like p
22	111.5	10.5	886	20 AAY98289	ppDE46 human duncel
23	111.5	10.5	886	21 AAB20639	Novel human diagno
24	111.5	10.5	901	22 ABG07070	Amino acid sequenc
25	111.5	10.5	1029	22 AAG65779	Amino acid sequenc
26	111.5	10.5	1124	22 AAG65777	Amino acid sequenc
27	111.5	10.5	1132	22 ANG65776	Amino acid sequenc
28	111.5	10.5	1132	22 AAG65778	Amino acid sequenc
29	111.5	10.5	1132	22 AAG65782	Drosophila melanog
30	110.5	10.4	531	22 ABG21540	WASP homolog prote
31	108.5	10.2	885	17 AAR99749	Amino acid sequenc
32	106.5	10.0	620	19 AAW72179	Amino acid sequenc
33	106.5	10.0	620	19 AAW72129	Amino acid sequenc
34	106.5	10.0	689	19 AAW72053	Amino acid sequenc
35	106.5	10.0	950	22 ABB71513	Human ataxin-2 like
36	105.5	9.9	574	22 AAM52322	Human polypeptide
37	105.5	9.9	574	22 AAG657370	Human protein SEQ
38	104.5	9.8	1081	21 AAB26240	Human N-methyl-D-a
39	104.5	9.8	1232	21 ABG26239	Human N-methyl-D-a
40	103.5	9.7	1040	22 ABG14734	Novel human diagno
41	103.5	9.7	1051	20 AAY29321	Human ataxin-2 like
42	102	9.6	321	22 AAM93743	Human protein sequ
43	102	9.6	321	22 AAB08090	Novel human diagno
44	101	9.5	230	22 AAB08090	Human protein SEQ
45	101	9.5	369	22 AAM79489	Human sapiens.

ALIGNMENTS

RESULT 1	AAW98158	standard; Protein: 198 AA.
ID	AAW98158	
XX		
AC	AAW98158;	
XX		
DT	05..JUL..1999	(first entry)
XX		
DE		Human Bcl-2 interacting mediator of cell death Bim-EL isoform.
XX		
KW		Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis; cell cycle; human; cancer; autoimmune disease; degenerative disease; therapy; contraceptive; splice variant; isoform.
XX		
PH		Key
FT		Binding-site
FT		/note= "dynein light chain binding region"
FT		92..100
FT		/note= "BH3"
FT		121..135
FT		/note= "hydrophobic region"

PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 XX
 PI Adams J., Cory S., Huang DCS, O'Connor L, O'Reilly L;
 PII Puthalakath H, Strasser A;
 XX
 DR WPI; 1999-244030/20.
 XX
 N-PSDB; AAX24997.
 XX
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer treatment
 XX
 PS Claim 15; Page 102-103; 145pp; English.
 XX
 CC The present sequence is the extra long form (EL) of human Bim, or Bcl-2 interacting mediator of cell death, a novel member of the Bcl-2 family that is capable of inducing cell death (apoptosis) and which acts as a 'death-ligand' for certain members of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the only Bcl-2 homology region which it encompasses is BH3. It is the only BH3-only protein for which splice variants exist. These result in the expression of a variety of isoforms, i.e. Bim-S, Bim-L and Bim-EL. cDNAs (see AAW98158) encoding human Bim-L and Bim-EL (see AAW98158) were obtained from embryo and liver cDNA libraries using mouse bim cDNA as probe. Murine Bim-S, Bim-L and Bim-EL isoforms are also provided (see AAW98154-56). Binding the dynein light chain was shown to regulate the pro-apoptotic activity of Bim. Bim-S, the splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that cannot bind, couple or otherwise associate with dynein light chain. The identification of Bim permits the identification and rational design of a range of products for use in therapy, diagnosis, antibody generation and involving modulation of physiological cell death. These therapeutic molecules may act as either antagonists or agonists of Bim's function and will be useful in cancer, autoimmune or degenerative disease therapy. Increased Bim expression or Bim activity is useful, e.g. for treatment or prophylaxis in conditions such as cancer and deletion of autoreactive lymphocytes in autoimmune disease. Decreased Bim expression of Bim activity is useful in regulation such as under cytotoxic prevention of cell death or degeneration such as under cytotoxic conditions during e.g. gamma-irradiation and chemotherapy or during HIV/AIDS or other viral infections, ischemia, myocardial infarction, hypoxia, degenerative diseases or for prolonging the survival of cells being transplanted for treatment of disease. Since Bim is expressed in germ cells, modulating Bim expression or Bim activity is useful, e.g. as a contraceptive or method of sterilization by preventing generation of fertile sperm.
 XX
 SO Sequence 198 AA;

Query Match 100.0%; Score 1065; DB 20; Length 198;
 Best Local Similarity 100.0%; Pred. No. 2 4e-93;
 Matches 198; Conservative 0; Mismatches 0; Gaps 0;

QY 1 MAKOPSDVSSCDEREGROLQAPRPPOLRGA
 Db 1 makqpsdsssecaregrg1qpaerppqrpqgaps1iqepqngpeqngedscphgsp 60
 QY 61 QGPLAPPASPGPFRATRSPFLFIMRSSLSSRSGGYESFDTDSAPMSCDSTSQTSP 120
 Db 61 qgplappaspgpfratrsfimrsslssrsgg1sfdtorspapmscdstsqtsp 120
 QY 121 CQAFNHYSAMSHRQAQADPAPMPEIWAQEELRIGDDEFNAYARRYFLNNYQAEDHPR 180
 Db 121 cqafnhysamshrqaaqadpapmpeiwaqeelrigddefnayarryflnnyqaedhpr 180
 QY 181 MVTLLRLRYIVRLYWRMH 198
 Db 181 mvrlrlryivrlvwrmh 198

RESULT 2
 AAW98156
 ID AAW98156 standard; Protein; 196 AA.
 XX
 AC AAW98156;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE Murine Bcl-2 interacting mediator of cell death Bim-EL isoform.
 XX
 KW Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; mouse; cancer; autoimmune disease;
 KW degenerative disease; therapy; contraceptive; splice variant;
 KW isoform.
 XX
 OS Mus musculus.
 XX
 FH Location/Qualifiers
 FT Binding-site 42-127
 /note= "dynein light chain binding region"
 FT Region 42-97
 /note= "present only in Bim-EL"
 FT Region 150-158
 /note= "BH3"
 FT Region 179-193
 /note= "hydrophobic region"
 XX
 PN W0914321-A1.
 XX
 PD 25-MAR-1999.
 XX
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 XX
 PR 17-SEP-1998; 98WO-AU00772.
 XX
 PR 24-SEP-1997; 97AU-0009373.
 XX
 PR 17-SEP-1997; 97AU-0009263.
 XX
 PR 1999-244030/20.
 DR N-PSDB; AAX24995.
 XX
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer treatment
 XX
 PS Claim 11; Page 98; 145pp; English.

CC The present sequence is the extra long form (EL) of murine Bim, or Bcl-2 interacting mediator of cell death, a novel member of the Bcl-2 family that is capable of inducing cell death (apoptosis) and which acts as a 'death-ligand' for certain members of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the only Bcl-2 homology region which it encompasses is BH3. It is the only BH3-only protein for which splice variants exist. These result in the expression of a variety of isoforms, i.e. Bim-S, Bim-L and Bim-EL (see AAW98154-56).
 CC cDNAs encoding these murine Bim isoforms (see AAW98153-95) were obtained from a T lymphoma cDNA library using human recombinant Bcl-2 protein. Human Bim-L and Bim-EL isoforms have also been identified (see AAW98157-58). Binding the dynein light chain was shown to regulate the pro-apoptotic activity of Bim. Bim-S, the splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain. The identification of Bim permits the identification and rational design of a range of products for use in therapy, diagnosis, antibody generation and involving modulation of physiological cell death. These therapeutic molecules may act as either antagonists or agonists of Bim's function and will be useful in cancer, autoimmune or degenerative disease therapy. Increased Bim expression or Bim activity is useful, e.g. for use in therapy, diagnosis, antibody generation and involving modulation of physiological cell death. These therapeutic molecules may act as either antagonists or agonists of Bim's function and will be useful in cancer, autoimmune or degenerative disease therapy. Increased Bim expression or Bim activity is useful, e.g. for

treatment or prophylaxis in conditions such as cancer and deletion of auto-reactive lymphocytes in autoimmune disease. Decreased Bim expression of Bim activity is useful in regulating inhibition or prevention of cell death or degeneration such as under cytotoxic conditions during e.g. gamma-irradiation and chemotherapy or during HIV/AIDS or other viral infections, ischemia, myocardial infarction, hypoxia, degenerative diseases or for prolonging the survival of cells being transplanted for treatment of disease. Since Bim is expressed in germ cells, modulating Bim expression or Bim activity is useful, e.g. as contraceptive or method of sterilization by preventing generation of fertile sperm.

Sequence	196 AA;	Query Match	85.3%	Score	908;	DB	20;	Length	196;
		Best Local Similarity	86.5%	pred.	No.	2e-78;			
		matches	173;	Conservative	10;	Mismatches	11;	Indels	6;
								Gaps	
1	MAKOPDSVSSCDBRGROLQDAERPQLRGAPTSLOTEPOGNPEGNHNCGEGDSCPHGSP	60							
1	makopdsvsscDBRGROLQDAERPQLRGAPTSLOTEPOGNPEGNHNCGEGDSCPHGSP	60							
1	makopdsvsscDBRGROLQDAERPQLRGAPTSLOTEPOGNPEGNHNCGEGDSCPHGSP	60							
61	QGFLAPPASPGFPATRSPRLPFEMR RSSLISRSSSSGYESFDTRSPAPMSCKSTQTSPSP	120							
57	qgflappaspgfpatsrlpifvrssllssssggfssfdlrsppmscdkstqtpssp	116							
121	CQAFNHYLSMASMROA - EPADMRPEIWIAQELRRIGDEFNAYAYARRFLNNYQAEDH	178							
117	cqafnhylsamas11gqeepledipeiraglrlrigdefnetyrrfandyreadch	176							
179	PRMVLRLRLLYIVRLIVWRMH	198							
177	prmvlrlrllyivrlivwrh	196							

RESULTS 3
498164
AAW98164 standard: Protein; 138 AA.

AAW98164;
05-JUL-1999 (first entry)
Human Bim-L mutant D51G.
Bim-L: Bcl-2 interacting mediator of cell death; apoptosis;
cell cycle; human cancer; autoimmune disease; therapy;
contraceptive; splice variant; isoform; mutant; dynein light chain;
cytotoxicity; agonist.

Key	Location/Qualifiers
Binding-site	42..71 'note= "dynein light chain binding region, present only in Bim-L, Bim-EL"
Misc-difference	51 'note= "Gly replaces Asp of wild-type Bim-L" 9..102 'note= "BH3" 123..137
Region	
Region	

WO9914321-A1.
25-MAR-1999.
17-SEP-1998;
24-SEP-1997;
17-SEP-1997;
98WO-AU000772.
97AU-0009373.
97AU-0009263.

XX SQ Sequence 138 AA;

Query	Match	65.4%	Score 696;	DB 20;	Length 138;
Best Local Similarity	69.7%	Pred. No. 1.3e-58;			
Matches 138; Conservative	0;	Mismatches 0;	Indels 60;	Gap	
Db	1	MAKQPSDVSSSECDEGRQLOPAERPPQLRPGAPTSLOTEPQQNPEGNHGGEDSCPQHGPSP			
Qy	1	QGPLAPPASPGPATRSPLFIMMRSSLLSSSGYFSDFDRSPAPMSCKDSTQTPSPPP			
Db	1	makqpsdvsssecdegrqlqpaerppqlrpqaptslqtepqqnpegnhggedscphgp			
Qy	61	QAFNHYLSAMASMRQAPADMPEIWAQELRIGDEFNAAYARRYFLNNYQAAEDHPR			
Db	42	l l l l l l l l l l l			
Qy	61	cqafnhylsamasmrqaspadmpeiwqaiqelrigdefnayarryflnnnyqaaedhpr			
Db	61	l l l l l l l l l l l			
Qy	181	MVILRLRYIVRLWRMH 198			
Db	121	mvilrlrlrvivrlwrmh 138			

RESULT	4
AAW98157	
ID	AAW98157 standard; protein: 138 AA.
XX	
AC	AAW98157:
XX	
DT	05-JUL-1999 (first entry)
XX	
DE	Human Bcl-2 interacting mediator of cell death Bim-L isoform.
XX	
KW	Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
XX	
KW	cell cycle; human; cancer; autoimmune disease;
KW	apoptosis; caspase-3; bcl-2; apoptosis; caspase-2; caspase-1; caspase-4;

KW	isoform.	Query	65.4%;	Score	696;	DB	20;	Length	138;
XX	Homo sapiens.	Best Local Similarity	69.7%;	Pred.	No. 1.8e-58;				
OS		Matches	138;	Mismatches	0;	Indels	60;	Gaps	1;
XX		Location/Qualifiers							
Key									
Binding-site	42..71	QY	1	MAKOPSDSSECDCREGROLQPAERPPQQLRGAPTSLSQTEPOQNGNEGNHCGDSCPHGSP	60				
FT	/note= "dynein light chain binding region, present	Db	1	makopsdsscdccregdqpaerppqqlrgaptslsqtep	41				
FT	only in Bim-L, Bim-EL"								
Region	92..100	QY	61	QGFLAPPASPSPGPATRSPPLFIMMRSSSLSSRSSSGYFSEFDTDSPPASCDKSTQTPSP	120				
FT	/note= "BHQ3"	Db	42	/dssppnscdkstqtp	60				
FT	121..135								
FT	/note= "hydrophobic region"								
XX	W09914321-A1.	QY	121	CQAFNHYLSAMASMRQAEPAADMPEIWIAQELRIGDEFENAYIARRVFLNNYQAAEDIPR	180				
XX	PD 25-MAR-1999.	Db	61	cqainhylsamasmrqaeapadmprpeiwiaqelrigrqdefenayarrvflnnyqaaedipr	120				
XX	PF 17-SEP-1998;	QY	181	MVILRLRLRYIVRLVWRMH	198				
XX	98W0-AU00772.	TD	AAW98165	standard; Protein;	138 AA.				
PR	24-SEP-1997;	XX							
PR	97AU-0009373.	AC	AAW98165;						
PR	17-SEP-1997;	DB	121	mvilrlrlryivrlvwrmh	138				
XX		RESULT	5						
PA	(HALL-) HALL INST MEDICAL RES WALTER & ELIZA.	AAW98165							
XX		ID	AAW98165						
PI	Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;	XX							
PI	Puthalakath H, Strasser A;	AC	AAW98165;						
XX		XX							
DR	WPI: 1999-244030/20.	DT	05-JUL-1999	(first entry)					
XX	N-PSDB; AAX24996.	XX							
PT	New isolated member of the Bcl-2 family, Bim used in, e.g. cancer	DE	Human Bim-L mutant S53P.						
PT	treatment	XX							
XX		KW	Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;						
PS	Claim 15; Page 100-101; 145pp; English.	XX	cell cycle; human; cancer; autoimmune disease; therapy;						
PS		XX	KW contraceptive; splice variant; isoform; mutant; dynein light chain;						
CC	The present sequence is the long form (L) of human Bim, or Bcl-2	XX	KW cytotoxicity; agonist.						
CC	interacting mediator of cell death, a novel member of the Bcl-2	XX	XX Homo sapiens.						
CC	family that is capable of inducing cell death (apoptosis) and which	XX	XX						
CC	acts as a 'death'-ligand, for certain members of the pro-survival	XX	XX						
CC	Bcl-2 family. Bim is a BH3-only protein, as the only Bcl-2 homology	XX	XX						
CC	region which it encompasses is BH3. It is the only BH3-only protein	XX	XX						
CC	for which splice variants exist. These result in the expression of	XX	XX						
CC	CC variants of isoforms, i.e. Bim-S, Bim-L and Bim-EL. cDNAs (see	XX	XX						
CC	AAZ996-97) encoding human Bim-L and Bim-EL (see AAW98158) were	XX	XX						
CC	obtained from embryo and liver cDNA libraries using mouse bim cDNA	XX	XX						
CC	as probe. Murine Bim-S, Bim-L and Bim-EL isoforms are also	XX	XX						
CC	provided (see AAW98154-56). Binding the dynein light chain was shown	XX	XX						
CC	to regulate the pro-apoptotic activity of Bim. Bim-S, the splice	XX	XX						
CC	variant which does not bind to dynein light chain, is a much more	XX	XX						
CC	potent killer than either Bim-L or Bim-EL. The invention provides	XX	XX						
CC	CC variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that	XX	XX						
CC	CC cannot bind, couple or otherwise associate with a dynein light	XX	XX						
CC	CC chain. The identification of Bim permits the identification and	XX	XX						
CC	CC rational design of a range of products for use in therapy,	XX	XX						
CC	CC diagnosis, antibody generation and involving modulation of	XX	XX						
CC	CC physiological cell death. These therapeutic molecules may act	XX	XX						
CC	CC as either antagonists or agonists of Bim's function and will be	XX	XX						
CC	CC useful in cancer, autoimmune or degenerative disease therapy.	XX	XX						
CC	CC Increased Bim expression or Bim activity is useful, e.g. for	XX	XX						
CC	CC treatment or prophylaxis in conditions such as cancer and deletion	XX	XX						
CC	CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim	XX	XX						
CC	CC expression of Bim activity is useful in regulating inhibition or	XX	XX						
CC	CC prevention of cell death or degeneration such as under cytotoxic	XX	XX						
CC	CC conditions during e.g. gamma-irradiation and chemotherapy or during	XX	XX						
CC	CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,	XX	XX						
CC	CC hypoxia, degenerative diseases or for prolonging the survival of	XX	XX						
CC	CC cells being transplanted for treatment of disease. Since Bim is	XX	XX						
CC	CC expressed in germ cells, modulating Bim expression or Bim activity	XX	XX						
CC	CC is useful, e.g. as a contraceptive or method of sterilization by	XX	XX						
CC	CC preventing generation of fertile sperm.	XX	XX						
Sequence	138 AA;	PS	97AU-0009373.						
Sequence	138 AA;	PR	97AU-0009263.						
Sequence	138 AA;	PD	17-SEP-1997.						
Sequence	138 AA;	XX	XX						
Sequence	138 AA;	XX	(HALL-) HALL INST MEDICAL RES WALTER & ELIZA.						
Sequence	138 AA;	XX	Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;						
Sequence	138 AA;	PI	Puthalakath H, Strasser A;						
Sequence	138 AA;	DR	WPI: 1999-244030/20.						
Sequence	138 AA;	PT	New isolated member of the Bcl-2 family, Bim used in, e.g. cancer						
Sequence	138 AA;	PT	treatment						
Sequence	138 AA;	XX	Claim 25; Page - ; 145pp; English.						
Sequence	138 AA;	XX	The present sequence is a S53P mutant of the long form (L) of						

human Bim, or Bcl-2 interacting mediator of cell death (see also AAW98157), a novel member of the Bcl-2 family. Native Bim-L is capable of inducing cell death (apoptosis) and acts as a 'death-ligand' for certain members of the pro-survival Bcl-2 family. Binding the dynein light chain regulates the pro-apoptotic activity of Bim. Bim-S (see AAW98154), a splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind couple or otherwise associate with a dynein light chain owing to amino acid addition, substitution and/or deletion within the region which binds the dynein light chain. The variants act as agonists of Bim function, interacting with anti-apoptotic molecules such as Bcl-2 to prevent their functional activity, thereby promoting apoptosis, and can be used e.g. in the treatment of cancer or to delete autoreactive lymphocytes in autoimmune diseases. Since Bim is expressed in germ cells, modulating its expression or activity may be useful as a contraceptive or method of sterilization by preventing generation of fertile sperm.

Note: The present sequence is not shown in the specification but is derived from the human Bim-L sequence given on page 100-101.

Sequence 138 AA;

Query Match	Best Local Similarity	Score	DB 20;	Length
Best Local Matches	64.9%	691;	Score 691;	138;
Local Similarity	69.28;	Pred. No. 5.4e-58;	DB 20;	Length 138;
Conservative	0;	Mismatches 1;	Length 138;	1.
Matches	137;	Indels 60;	Gaps	
1	MAKQPSDVSECDREGROLQPAERPQLRPGAPTSLOTEPGQNPEGNHIGEGDSCPQHGSP	60		
1	makqpsdvsecdregrolqpaerpqlrpgaptsloqpegnhigegdscphgsp	60		
1	makqpsdvsecdregrolqpaerppqirgaptqlqeqp	60	-----	41
61	QGPLAPPASPGPATRSPFLFIMPRSSLURSSSGYFSSTDTDRSPAMSCDKSTQTPSPPP	120		
61	qgplappaspgpatrspflfimprsslursssgyfsstdtdrspamscdkstqtpsp	120		
42	-----	60	-----	
121	COAFNHYLSAMSMRQAEPADMRPEIWIQAELRRLIGDEFNAYARRVFLNNYQAAEDHPR	180		
121	coafnhylsamsmrqae padmrpeiwiqaelrrrigdefnayarrvflnnynqaedhpr	180		
61	cqafnhylsamsmrqae padmrpeiwiqaelrrrigdefnayarrvflnnynqaedhpr	120		
181	MVILRLIIRRIVRLYWRMH	198		
181	mvilrlirrivyrlivrlwrmh	198		
121	mvilrlirrivyrlivrlwrmh	138		

Homo sapiens.		
Synthetic.		
	Key	Location/Qualifiers
	Binding-site	42..71 /note- "dynein light chain binding region, pre- only in Bim-L, Bim-EL"
	Misc-difference	54 /note- "Ala replaces Thr of wild-type Bim-L"
	Region	94..102 /note- "BH3" 123..137
	Region	

FT /note= "hydrophobic region"
XX W09914321-A1.
PN XX
XX 25 -MAR-1999.
PD XX
XX PF 17 -SEP-1998; 98WO-AU00772.
XX PR 24 -SEP-1997; 97AU-0009373.
XX PR 17 -SEP-1997; 97AU-0009263.
XX PA (HALL -) HALL INST MEDICAL RES WALTER & ELIZ.
XX XX
PT Adams J, Cory S, Huang DCS, O'Connor L,
PI Puthalakath H, Strasser A;
XX DR WPT; 1999-244030/20.
XX PT New isolated member of the Bcl-2 family, Bcl-2
PT treated
XX

תְּמִימָנָה (תְּמִימָנָה) בְּמִינָה (בְּמִינָה)

human Bim, or Bcl-2 interacting mediator of cell death (see also AW98157), a novel member of the Bcl-2 family. Native Bim-L is capable of inducing cell death (apoptosis) and acts as a death-ligand, for certain members of the pro-survival Bcl-2 family. Binding the dynein light chain regulates the pro-apoptotic activity of Bim. Bim-S (see AW98154), a splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AW98155-68) of murine and human Bim-L and Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain owing to amino acid addition, substitution and/or deletion within the region which binds the dynein light chain. The variants act as agonists of Bim function, interacting with anti-apoptotic molecules such as Bcl-2 to prevent their functional activity, thereby promoting apoptosis, and can be used e.g. in the treatment of cancer or to delete auto-reactive lymphocytes in autoimmune diseases. Since Bim is expressed in germ cells, modulating its expression or activity may be useful as a contraceptive or method of sterilization by preventing generation of fertile sperm. Note. The present sequence is not shown in the specification but is derived from the human Bim-L sequence given on page 100-101.

XX	AC	AAW98168;	Query	Match	64.9%;	Score	691;	DB	20;	Length	138;
XX	AC	05-JUL-1999	Best Local Matches	69.2%;	Pred. No.	5.4e-58;	Indels	60;	Gaps	1;	
DT	DE	(first entry)	Matches	137;	Conservative	1;	Mismatches	0;			
XX	Human	Bim-L mutant N65S.									
XX	Bim-L; Bcl-2 interacting mediator of cell death; apoptosis; cell cycle; human; cancer; autoimmune disease; therapy; contraceptive; splice variant; isoform; mutant; dynein light chain; cytotoxicity; agonist.	1	MAKOPSDVSECDREGRQLOPAERPPQLRGPAPTSLOTEPQGNPEGNIGEGDSCPHGSP	60							
XX	KW		1 makopsdvssecdregrqlopaerppqlrgpaptsloqeq-----	41							
XX	KW		61 QGPIAPPASPGPFATRSPLFIFMRSSSLRSRSSCGYFSFDTDRSPAPMCDKSTQTPSP	120							
XX	KW		42 -----drspapmcdksrgtpp	60							
XX	KW		61 cqafnhylsamasmrqapeadmrpeiwiaqelrrtgdefnayarrvflnnqyaaedhpr	180							
XX	KW		61 cqafnhylsamasmrqapeadmrpeiwiaqelrigdeinayarrvflnnqyaaedhpr	120							
XX	OS	Homo sapiens.									
XX	OS	Synthetic.									
FH	Key										
FT	Binding-site										
FT		/note= "dynein light chain binding region, present only in Bim-L, Bim-EL"									
FT	Misc-difference	65									
FT		/note= "Ser replaces Asn of wild-type Bim-L"									
FT	Region	94..102	RESULT	8							
FT		/note= "BH3"	ID	AAW98167	standard;	protein;	138	AA.			
FT	Region	123..137	XX								
FT		/note= "hydrophobic region"	AC	AAW98167;							
XX	W09914321-A1.		XX								
XX	PD	25-MAR-1999.	XX								
XX	PD		DE								
XX	PF	17-SEP-1998;	Human	Bim-L mutant T54 I.							
XX	PF	98W0-AU00772.	XX								
XX	PR	24-SEP-1997;	KW	Bim-L; Bcl-2 interacting mediator of cell death; apoptosis; cell cycle; human; cancer; autoimmune disease; therapy;							
XX	PR	97AU-0009373.	KW	contraceptive; splice variant; isoform; mutant; dynein light chain; cytotoxicity; agonist.							
XX	PA	17-SEP-1997;	KW								
PA	(HALL-) HALL INST MEDICAL RES WALTER & ELIZA.	XX	XX								
XX	PA	(HALL-) HALL INST MEDICAL RES WALTER & ELIZA.	OS	Homo sapiens.							
PI	Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;	OS	OS	Synthetic.							
PI	Puthalakath H, Strasser A;	XX	XX								
XX	DR	WPI; 1999-244030/20.	FH	Location/Qualifiers							
XX	DR		FT	Binding-site	42..71	"dynein light chain binding region, present only in Bim-L, Bim-EL."					
XX	PT	New isolated member of the Bcl-2 family, Bim used in, e.g. cancer treatment	FT	Misc-difference	54	/note= "Ile replaces Thr of wild-type Bim-L"					
XX	PS	Claim 25; Page - ; 145pp; English.	FT	Region	94..102	/note= "BH3"					
XX	CC	The present sequence is a N65S mutant of the long form (L) of human Bim, or Bcl-2 interacting mediator of cell death (see also AAW98157), a novel member of the Bcl-2 family. Native Bim-L is capable of inducing cell death (apoptosis) and acts as a 'death-ligand', for certain members of the pro-survival Bcl-2 family. Binding the dynein light chain regulates the pro-apoptotic activity of Bim. Bim-S (see AAW98154), a splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain owing to, amino acid addition, substitution and/or deletion within the region which binds the dynein light chain. The variants act as agonists of Bim function, interacting with anti-apoptotic molecules such as Bcl-2 to prevent their functional activity, thereby promoting apoptosis, and can be used e.g. in the treatment of cancer or to delete auto-reactive lymphocytes in autoimmune diseases. Since Bim is expressed in germ cells, modulating its expression or activity may be useful as a contraceptive or method of sterilization by preventing generation of fertile sperm.	XX	FT	Region	123..137	/note= "hydrophobic region"				
CC	CC	Note. The present sequence is not shown in the specification but is derived from the human Bim L sequence given on page 100-101.	XX	FT	Region	25-MAR-1999;	97AU-0009373.				
CC	CC	Sequence 138 AA;	XX	FT	Region	17-SEP-1998;	98W0-AU00772.				
CC	CC	PS Claim 25; Page - ; 145pp; English.	XX	PA	(HALL-) HALL INST MEDICAL RES WALTER & ELIZA.						
CC	CC	XX	PA	Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L; Puthalakath H, Strasser A;							
CC	CC	XX	PR	24-SEP-1997;	97AU-0009373.						
CC	CC	XX	PR	17-SEP-1997;	97AU-0009263.						
CC	CC	XX	DR	WPI; 1999-244030/20.							
CC	CC	XX	PA	New isolated member of the Bcl-2 family, Bim used in, e.g. cancer treatment							
CC	CC	XX	PT	The present sequence is a T54I mutant of the long form (L) of human Bim, or Bcl-2 interacting mediator of cell death (see also							

AAW98157), a novel member of the Bcl-2 family. Native Bim-L is capable of inducing cell death (apoptosis) and acts as a 'death-ligand' for certain members of the pro-survival Bcl-2 family. Binding the dynein light chain regulates the pro-apoptotic activity of Bim. Bim-S (see AAW8154), a splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind couple or otherwise associate with a dynein light chain owing to amino acid addition, substitution and/or deletion within the region which binds the dynein light chain. The variants act as agonists of Bim function interacting with anti-apoptotic molecules such as Bcl-2 to prevent their functional activity, thereby promoting apoptosis, and can be used e.g. in the treatment of cancer or to delete autoreactive lymphocytes in autoimmune diseases. Since Bim is expressed in germ cells, modulating its expression or activity may be useful as a contraceptive or method of sterilization by preventing generation of fertile sperm. Note. The present sequence is not shown in the specification but is derived from the human Bim-L sequence given on page 100-101.

Query	Match	Score	DB	Length
Query	1	64.8%	690	20
Best Local Similarity		69.2%	Pred. No.	138;
Matches	137;	Conservative	6.7e-58;	
		Matches	0;	Mismatches
			1;	Indels
			60;	Gaps
Qy	1	MAKOPSDVSSECDRGQLQPAERPQLRQGAPTSLOTEPGNPECNHGEGDSCPHGSP	60	
Ddb	1	makopdsvssecdrqqlqpaerpqlrqgaptslqterg-----	41	
Qy	61	QGPIAPPASPQPFATRSPLFIFMRRSSLLSRSSGGYFSFTIDRSRAPMSCDKSTQTQSPSP	120	
Ddb	42	-----drspampsodks14tpssp 60	120	
Qy	121	CQAFNHYLSAMASMRQAEPADMRPEIWIQAEQELRRIGDEFNAYAYARRFLNNYQAAEDHPR	180	
Ddb	61	cqafnhylsamasmrqaeypadmrpeiwiaqelrrigdefnayayarrflnnyqaaedhpr	120	
Qy	181	MVILRLRLTVRLWRMH	198	

Mus	<i>mus musculus</i>	
OS		
XX		
Key		Location/Qualifiers
Brinding-site		42..71
FFT		/note= "dynein light chain binding region, pre-
FFT		only in Bim-L, Bim-EL"
Region		94..102
FFT		/note= "BH3"
Region		123..137
FFT		

Db	61	cgafnhylsamasirsgsgeedlrpeiriaqellrigdefnetyrrvfandyreadh 120
Qy	179	PRMVVILRLRYVTRLYWWRMH 198 : : : : 121 pgmviqllriffrivrrh 140
Db	AAW98160	standard; Protein; 140 AA.
	XX	
	XX	
	XX	05-JUL-1999 (first entry)
	XX	
	DE	Murine Bim-L mutant S53P.
	XX	
	KW	Bim-L; Bcl-2 interacting mediator of cell death; apoptosis; cell cycle; mouse; cancer; autoimmune disease; therapy; contraceptive; splice variant; isoform; mutant; dynein light chain; cytotoxicity; agonist.
	XX	
	OS	Mus musculus.
	OS	Synthetic.
	XX	
	FH	Key
	FT	Binding-site
	FT	42..71
	FT	/note= "dynein light chain binding region, Present only in Bim-L, Bim-EL"
	FT	
	Misc-difference	53
	FT	/note= "pro replaces Ser of wild-type Bim-L"
	Region	94..102
	FT	/note= "BH3"
	FT	123..137
	FT	/note= "hydrophobic region"
	XX	
	PN	W0991321-A1.
	XX	
	PD	25-MAR-1999.
	XX	
	PF	17-SEP-1998;
	XX	98WO-AU00772.
	PR	24-SEP-1997;
	PR	97AU-0009373.
	XX	17-SEP-1997;
	PR	97AU-0009263.
	PA	(HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
	XX	
	PI	Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
	PI	Putthakath H, Strasser A;
	XX	
	DR	WPI; 1999-244030/20.
	PT	New isolated member of the Bcl-2 family. Bim used in, e.g. cancer treatment
	XX	
	PS	Claim 25; Page ; 145pp; English.
	XX	
CC	CC	The present sequence is a S53P mutant of the long form (L) of murine Bim, or Bcl-2 interacting mediator of cell death (see also AW98159), a novel member of the Bcl-2 family. Native Bim-L is capable of inducing cell death (apoptosis) and acts as a 'death ligand', for certain members of the pro-survival Bcl-2 family. Binding the dynein light chain regulates the pro-apoptotic activity of Bim. Bim-S (see AW98154), a splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AW98159-68) of murine and human Bim-L and Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain owing to amino acid addition, substitution and/or deletion within the region which binds the dynein light chain. The variants act as agonists of Bim function, interacting with anti-apoptotic molecules such as Bcl-2 to prevent their functional activity, thereby promoting to apoptosis, and can be used e.g. in the treatment of cancer or to

XX DR WP1; 1999-244030/20.
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer treatment
 XX PT
 PS Claim 25: Page -; 145pp; English.
 XX The present sequence is a T54A mutant of the long form (L) of murine Bim, or Bcl-2 interacting mediator of cell death (see also AAW98155), a novel member of the Bcl-2 family. Native Bim-L is capable of inducing cell death (apoptosis) and acts as a 'death-ligand', for certain members of the pro-survival Bcl-2 family. Binding the dynein light chain regulates the pro-apoptotic activity of Bim. Bim-S (see AAW98154), a splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain owing to amino acid addition, substitution and/or deletion within the region which binds the dynein light chain. The variants act as agonists of Bim function, interacting with anti-apoptotic molecules such as Bcl-2 to prevent their functional activity, thereby promoting apoptosis, and can be used e.g. in the treatment of cancer or to delete autoreactive lymphocytes in autoimmune diseases. Since Bim is expressed in germ cells, modulating its expression or activity may be useful as a contraceptive or method of sterilization by preventing generation of fertile sperm.
 CC Note. The present sequence is not shown in the specification but is derived from the murine Bim-L sequence given on page 95-96.
 XX Sequence 140 AA;
 SQ

Query Match 54.8%; Score 584; DB 20; Length 140;
 Best Local Similarity 59.5%; Pred. No. 8.1e-48;
 Matches 119; Conservative 8; Mismatches 11; Indels 62; Gaps 2;

QY 1 MAKQPSDVSSCDREGQLQPAERPPOLRPGAPTSLOTEPPQNPEGNHGGBDSCPQHGSP 60
 Db 1 makqpsdvsscdregqlqpaerppolrpgaptslo... 41
 QY 61 QGQPLAPPASPGPFATRSPLFIFMRSSLSSSGYFSEDTDRSPAPMSCDKSTQTPSPPP 120
 Db 42 ... 60
 QY 121 COAFNHYLSAMASMROA-EPADMREPEIWIQAELRRIGDEFNAYYARRVFLNNYQAEDH 178
 Db 61 caafnhylsamasmroa-epadmrepeiwiaqlrrigdefnayyarrvflnnyqaedh 120
 QY 179 PRMVILLLRYIVRLWYRMH 198
 Db 121 prmvilllryivrlwyrmh 140

RESULT 12
 ID AAW98163 standard; Protein: 140 AA.
 XX AAW98163;
 AC AAW98163;
 DT 05-JUL-1999 (first entry)
 XX Murine Bim-L mutant N55S.
 KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; mouse; cancer; autoimmune disease; therapy;
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;
 KW cytotoxicity; agonist.
 XX Mus musculus.
 OS Synthetic.
 XX

FH Key Binding-site Location/Qualifiers
 FT FT 42..71 /note= "dynein light chain binding region, present only in Bim-L, Bim-EL"
 FT FT 65 /note= "Ser replaces Asn of wild-type Bim-L"
 FT FT 94..102 /note= "BH3"
 FT FT 123..137 /note= "hydrophobic region"
 XX WO9914321-A1.
 PN PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 PD 25 MAR-1999.
 XX PF 17 SEP-1998; 98WO-A000772.
 XX PR 24 SEP-1997; 97AU-0009373.
 PR 17 SEP-1997; 97AU-0009263.
 XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;
 DR WPI; 1999-244030/20.
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer treatment
 XX PS Claim 25: Page -; 145pp; English.
 CC The present sequence is a N65S mutant of the long form (L) of murine Bim, or Bcl-2 interacting mediator of cell death (see also AAW98155), a novel member of the Bcl-2 family. Native Bim-L is capable of inducing cell death (apoptosis) and acts as a 'death-ligand', for certain members of the pro-survival Bcl-2 family. Binding the dynein light chain regulates the pro-apoptotic activity of Bim. Bim-S (see AAW98154), a splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain owing to amino acid addition, substitution and/or deletion within the region which binds the dynein light chain. The variants act as agonists of Bim function, interacting with anti-apoptotic molecules such as Bcl-2 to prevent their functional activity, thereby promoting apoptosis, and can be used e.g. in the treatment of cancer or to delete autoreactive lymphocytes in autoimmune diseases. Since Bim is expressed in germ cells, modulating its expression or activity may be useful as a contraceptive or method of sterilization by preventing generation of fertile sperm.
 CC Note. The present sequence is not shown in the specification but is derived from the murine Bim-L sequence given on page 95-96.
 XX Sequence 140 AA;
 SQ

Query Match 54.8%; Score 584; DB 20; Length 140;
 Best Local Similarity 59.5%; Pred. No. 8.1e-48;
 Matches 119; Conservative 9; Mismatches 10; Indels 62; Gaps 2;

QY 1 MAKQPSDVSSCDREGQLQPAERPPOLRPGAPTSLOTEPPQNPEGNHGGBDSCPQHGSP 60
 Db 1 makqpsdvsscdregqlqpaerppolrpgaptslo... 41
 QY 61 QGQPLAPPASPGPFATRSPLFIFMRSSLSSSGYFSEDTDRSPAPMSCDKSTQTPSPPP 120
 Db 42 ... 60
 QY 121 COAFNHYLSAMASMROA-EPADMREPEIWIQAELRRIGDEFNAYYARRVFLNNYQAEDH 178
 Db 61 caafnhylsamasmroa-epadmrepeiwiaqlrrigdefnayyarrvflnnyqaedh 120
 QY 179 PRMVILLLRYIVRLWYRMH 198
 Db 121 prmvilllryivrlwyrmh 140

Query Match 54.8%; Score 584; DB 20; Length 140;
 Best Local Similarity 59.5%; Pred. No. 8.1e-48;
 Matches 119; Conservative 9; Mismatches 10; Indels 62; Gaps 2;

QY 1 MAKQPSDVSSCDREGQLQPAERPPOLRPGAPTSLOTEPPQNPEGNHGGBDSCPQHGSP 60
 Db 1 makqpsdvsscdregqlqpaerppolrpgaptslo... 41
 QY 61 QGQPLAPPASPGPFATRSPLFIFMRSSLSSSGYFSEDTDRSPAPMSCDKSTQTPSPPP 120
 Db 42 ... 60

QY 121 CQAFNHYLSAMASMROA-EPADMREPEIWIQAELRRIGDEFNAYYARRVFLNNYQAEDH 178
 Db 61 cqafnhylsamasmroa-epadmrepeiwiaqlrrigdefnayyarrvflnnyqaedh 120

Query	Match	54	6%	Score	582;	DB	20;	Length	140;
Best	Local Similarity	59	5%	Pred.	No.	1	3e-47;		
Matches	119;	Conservative	8;	Missmatches	11;	Indels	62;	Gaps	2;
1	MAKQPSDVSSECDRSGQLOPAERPPQLRGPAPTSLSQTEPQGNPEGNIGEGDSCPAPSSP	60							
1	makqpsdvssecdregqqlqpaerppqlrgpaptslsqtepq-----	60							
41	QGPLAPPASPGPFATRSPFLFIMRSSLILSRSSGGYFSPFDTDRSPAPMSCDKSTOTPSPP	120							
41	qgplappaspgpfatrspflfimrsssllsrssggyfspfdtdrsppapmscdkstotpspp	120							
42	-----	-----							
42	-drspapnscgkstqtpsp	60							
121	CQAFNHYLSAMASMRQA -EPADMREPIWIAQELARRIGDEFNAYARRYFLNNYQAAEDH	178							
121	cqafnhylsamasirqssepedirpeirajaeirrigdefnetyrrfandyreaedh	178							
179	PRMVLRLRFLYIVRLWWRMH	198							
179	prmvlrlrlfylivrlwwrmh	198							
121	PQMV1QLIIRIFRIVWRH	140							
121	pqmv1qliriifrivwrh	140							

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 09:57:22 ; Search time 13.18 Seconds (without alignments)

366,940 Million cell updates/sec

Title: US-09-508-832-10

Perfect score: 1065

Sequence: 1 MAKQPSDVSSECDREGQLQ.....PRMVILRLRLYVRLVWRMH 198

Scoring table: BLOSUM62

Gapext 0.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/2/iaa/PCMB_COMB.pep:*

6: /cgn2_6/ptodata/2/iaa/backfileseqs.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111.5	10.5	886	2 US-08-474-379C-65	Sequence 65, Appl
2	111.5	10.5	886	3 US-09-16-249A-65	Sequence 65, Appl
3	111.5	10.5	886	3 US-08-206-188B-65	Sequence 65, Appl
4	108.5	10.2	885	2 US-08-577-492-33	Sequence 33, Appl
5	108.5	10.2	885	4 US-09-079-630-33	Sequence 33, Appl
6	99	9.3	228	4 US-09-219-849-38	Sequence 4, Appl
7	99	9.3	1057	3 US-08-931-820-4	Sequence 4, Appl
8	99	9.3	1078	3 US-08-96-825-21	Sequence 21, Appl
9	99	9.3	1078	4 US-09-573-21	Sequence 21, Appl
10	99	9.3	1078	4 US-09-548-608-21	Sequence 21, Appl
11	94.5	8.9	280	4 US-09-247-155-178	Sequence 178, Appl
12	94.5	8.9	1185	4 US-09-041-886-23	Sequence 23, Appl
13	94	8.8	580	3 US-08-906-865-1	Sequence 6, Appl
14	94	8.8	627	2 US-08-466-589-6	Sequence 6, Appl
15	94	8.8	627	3 US-08-467-574-6	Sequence 6, Appl
16	94	8.8	627	4 US-09-140-804-4	Sequence 6, Appl
17	94	8.8	627	4 US-09-217-345-6	Sequence 6, Appl
18	93.5	8.8	434	2 US-08-710-249-4	Sequence 4, Appl
19	93.5	8.8	434	4 US-09-220-157A-4	Sequence 4, Appl
20	93.5	8.8	902	1 US-08-336-479B-6	Sequence 6, Appl
21	93.5	8.8	902	1 US-08-818-823-6	Sequence 6, Appl
22	92.5	8.7	245	4 US-09-140-804-4	Sequence 4, Appl
23	92.5	8.7	829	1 US-08-642-255-132	Sequence 132, Appl
24	92.5	8.7	829	1 US-08-337-633A-53	Sequence 53, Appl
25	92	8.6	761	2 US-08-237A-84	Sequence 84, Appl
26	92	8.6	762	1 US-08-642-255-114	Sequence 114, Appl
27	92	8.6	762	1 US-08-642-255-120	Sequence 120, App

ALIGNMENTS

RESULT 1

US-08-474-379C-65

; Sequence 65, Application US/08474379C

; Patent No. 5977305

; GENERAL INFORMATION:

; APPLICANT: Wigler, Michael H.

; ATTORNEY: Colicelli, John J.

; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED

; TITLE OF INVENTION: PROCESSES

; NUMBER OF SEQUENCES: 88

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 South Wacker Drive/6300 Sears Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/474-379C

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 07/511,715

; FILING DATE: 20-APR-1990

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/205,188

; FILING DATE: 01-MAR-1994

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 07/680,352

; FILING DATE: 19-APR-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Cough, David W.

; REGISTRATION NUMBER: 36,107

; REFERENCE/DOCKET NUMBER: 27886/32771

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 474-6300

; TELEFAX: (312) 474-0448

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 986 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-474-379C-65

Query Match 10.5%; Score 111.5; DB 2; Length 886;
 Best Local Similarity 27.5%; Pred. No. 0.0046; Indels 39; Gaps 7;
 Matches 47; Conservative 18; Mismatches 67; Del 11; Insert 11;

QY 6 SDVSSECDCREGQIQLQPAERPPQL---RPGAPTSLSQTEPQ---GNPEGNHC-----50
 Db 49 SDSERAERERQHPPIERADMTSDRPLGLTRTRMSHPLSSFHGTGCGGAGGSSRRF 108

QY 51 EGDSCPHGPSP-QGLPAPPSPGPATREPLFIMRSSLSSGIGSSEFDTRSPAPMS 109
 Db 109 EAENGPTSPGRSPLDQASGP-LVLHAGAATSQRESFLYRSDY---DMSPKTMS 162

QY 110 CDKSTQTPSPCQAFNHYLSAMASMRQAEPADMRPEIWAQELRRIGDEFN 160
 Db 163 RN-----SSVTEAHAEDLIVTPFAQVLASLRSVRNSFS 196

RESULT 3
 US-08-206-188B-65
 Sequence 65, Application US/08206188B
 Patent No. 610025
 GENERAL INFORMATION:
 APPLICANT: Wigler, Michael H.
 APPLICANT: Colicelli, John J.
 TITLE OF INVENTION: Cloning by Complementation and Related
 NUMBER OF SEQUENCES: 84
 NUMBER OF SEQUENCES: 84
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/206,188B
 FILING DATE: 01-MAR-1994
 CLASSIFICATION: 435
 PRIORITY DATA:
 APPLICATION NUMBER: US 07/511,715
 FILING DATE: 20-APR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Clough, David W.
 REGISTRATION NUMBER: 36107
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 65:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 886 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-206-188B-65

Query Match 10.5%; Score 111.5; DB 3; Length 886;
 Best Local Similarity 27.5%; Pred. No. 0.0046; Indels 39; Gaps 7;
 Matches 47; Conservative 18; Mismatches 67; Del 11; Insert 11;

QY 6 SDVSSECDCREGQIQLQPAERPPQL---RPGAPTSLSQTEPQ---GNPEGNHC-----50
 Db 49 SDSERAERERQHPPIERADMTSDRPLGLTRTRMSHPLSSFHGTGCGGAGGSSRRF 108

QY 51 EGDSCPHGPSP-QGLPAPPSPGPATREPLFIMRSSLSSGIGSSEFDTRSPAPMS 109
 Db 163 RN-----SSVTEAHAEDLIVTPFAQVLASLRSVRNSFS 196

RESULT 4
 US-08-577-492-33
 Sequence 33, Application US/08577492
 Patent No. 5851784

GENERAL INFORMATION:

APPLICANT: Owens, Raymond John
Perry, Martin John

APPLICANT: Lamb, Simon Mark

TITLE OF INVENTION: HUMAN PHOSPHODIESTERASE TYPE IVC, AND ITS PRODUCTION AND USE

NUMBER OF SEQUENCES: 40

ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & NO. 5851784ris

STREET: One Liberty Place, 46th floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/577,492

FILING DATE: 22-DEC-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9426227.6

FILING DATE: 23-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9512996.1

CURRENT APPLICATION DATA:

FILING DATE: 26-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Cherry, David A.

REGISTRATION NUMBER: 35,099

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100

TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 885 amino acids

TYPE: amino acid

TOPOLOGY: unknown

US-08-577-492-33

Query Match 10 2%; Score 108 5; DB 2; Length 885;
Best Local Similarity 27.1%; Pred. No. 0.0092; Mismatches 69; Indels 37; Gaps 54

Matches 46; Conservative 18; Mismatches 69; Indels 37; Gaps 54

Qy 6 SDVSSCDBREGQLOQPAERPPQL ---RPGAPTSIQLQEP ---QCNPEGHGGEDS --- 54
Db 49 SDSERAEREROPRPIERADMTSDREGLRTMRMNPSSFHGTGTGGAGGSSRRF 108

Qy 55 -CPIGSPOQG -PLAPPASGP -FTRSPSLFIMMRSSLLRSSSGGYFSFSDTDRPAPMSC 110
Db 109 EAERSSRSPSLPMTSSPGLMVLHAGATTSQRPRESFYRSBDY ---DRSPKTMSPR 166

Qy 111 DKSTQTPSPCOAHPNLSAMASMRQAEPADEPEIWAQELRIGDEFN 160
Db 164 N-----SSVTSAHAEDLIVTPEAQVLASLRSVRSNFS 196

RESULT 5
US-09-079-630-33
; Sequence 33 Application US/09079630
; Patent No. 6291199

GENERAL INFORMATION:

APPLICANT: Owens, Raymond John
Perry, Martin John

APPLICANT: Lamb, Simon Mark

TITLE OF INVENTION: HUMAN PHOSPHODIESTERASE TYPE IVC, AND ITS PRODUCTION AND USE

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & NO. 6291199ris

STREET: One Liberty Place, 46th floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Wordperfect 6.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/079,630
 FILING DATE: 07-09-1998
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/577,492
 FILING DATE: 22-DEC-1995
 APPLICATION NUMBER: GB 9426227.6
 FILING DATE: 23-DEC-1994
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: GB 9512996.1
 FILING DATE: 26-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Cherry, David A.
 REGISTRATION NUMBER: 35,099
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3100
 TELEFAX: (215) 568-4439
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 885 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 US-09-079-630-33

Query Match 10.2%; Score 108.5
 Best Local Similarity 27.1%; Prcd. No. 0
 Matches 46; Conservative 18; Mismatch 0

Qy	6	SDVSSECDRREGRLQPAERPPOL---	-RPGAPTSLSA
Db	49	SDAEEAERERQPHPIERADMTDSRPLGLRTRI	
Qy	55	-CPHGSPOG -PLAPPASPGP-FATRSPLFIMMR	
Db	109	EAENGTSAGRSPLDPMSPSPGLVLYHAGAATSRRI	
Qy	111	DKSTQTPSPPCQAFNHYLSAMASMRQAPADMRPE	
Db	164	N-----SSVTESEAHDLELIVTPPE	

RESULT 6
 US-09-219-849-38
 Sequence 38, Application US/09/219849
 Patent No. 6150081
 GENERAL INFORMATION:
 APPLICANT: VAN HEERDE, GEORGE V.
 APPLICANT: VAN RIJN, ALEXIS C.
 APPLICANT: BOUWSTRA, JAN B.
 APPLICANT: DE WOLF, FREDERIK A.
 APPLICANT: MOOBROEK, ANDREAS
 APPLICANT: WENTEN, MARC W.T.
 APPLICANT: WIND, RICHELE D.
 APPLICANT: VAN DEN BOSCH, TANJA J.
 TITLE OF INVENTION: SILVER HALIDE EMULSION
 TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHY
 TITLE OF INVENTION: PREPARATION THEREOF
 FILE REFERENCE: 2728-2
 CURRENT APPLICATION NUMBER: US/09/219,849
 CURRENT FILING DATE: 1998-12-23
 NUMBER OF SEQ ID NOS: 50
 CURRENT FILING DATE: 1998-12-23

US-09-548-608-21 Sequence 21, Application US/09548608
 Patient No. 6355442

GENERAL INFORMATION:

APPLICANT: Qvist, Per
 BONDÉ, Martin
 ADDRESS: Bondé, Martin
 TITLE OF INVENTION: A Method for Assaying Collagen Fragments
 in Body Fluids, A Test Kit and Means for Carrying Out the
 Method and Use of the Method to Diagnose the Presence of
 Disorders Associated with the Metabolism of
 Collagen

TITLE OF INVENTION: Method for Assaying Collagen Fragments
 in Body Fluids, A Test Kit and Means for Carrying Out the
 Method and Use of the Method to Diagnose the Presence of
 Disorders Associated with the Metabolism of
 Collagen

TITLE OF INVENTION: Method for Assaying Collagen Fragments
 in Body Fluids, A Test Kit and Means for Carrying Out the
 Method and Use of the Method to Diagnose the Presence of
 Disorders Associated with the Metabolism of
 Collagen

TITLE OF INVENTION: Method for Assaying Collagen Fragments
 in Body Fluids, A Test Kit and Means for Carrying Out the
 Method and Use of the Method to Diagnose the Presence of
 Disorders Associated with the Metabolism of
 Collagen

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby & Darby PC
 STREET: 805 Third Avenue
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/548, 608
 FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/187, 319
 FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Goror is, Adda C.
 REGISTRATION NUMBER: 29, 714
 REFERENCE/DOCKET NUMBER: 4305/08701
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-527-7700
 TELEX: 236687

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 1078 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 IMMEDIATE SOURCE:
 CLONE: COLLAGEN ALPHA 1 (III)

US-09-548-608-21

C-terminus		N-terminus		C-terminus		N-terminus	
COLLAGEN ALPHA 1 (III)	-09-570-573-21	COLLAGEN ALPHA 1 (III)	-09-570-573-21	COLLAGEN ALPHA 1 (III)	-09-570-573-21	COLLAGEN ALPHA 1 (III)	-09-570-573-21
Query Match	9.3%	Score 99;	DB 4;	Length 1078;	Query Match	9.3%	Score 99;
Best Local Similarity	27.8%	Pred. No.	0.11;		Best Local Similarity	27.8%	Pred. No.
Matches	37;	Conservative	7;	Mismatches	45;	Matches	37;
Indels	44;	Gaps	5;	Indels	45;	Indels	44;
5	PSDYSSECDREGROLQPAERPPOLRPGAPTSLOQT-EP--QGNP-----EGNHGEG	52	Qy	5	PSDYSSECDREGROLQPAERPPOLRPGAPTSLOQT-EP--QGNP-----EGNHGEG	52	
826	PQGVKGESKPGANGLISGERGPPGQLPGLAGTAGEPGRDGNPGSDGLPGRDGKKG	885	Db	826	PQGVKGESKPGANGLISGERGPPGQLPGLAGTAGEPGRDGNPGSDGLPGRDGKKG	885	
53	DSCPNGSPQGPLAP--PASPGPFATRSPLFIMMRSSSLISRSSSSGYFSFTDRSPAPMSC	110	Qy	53	DSCPNGSPQGPLAP--PASPGPFATRSPLFIMMRSSSLISRSSSSGYFSFTDRSPAPMSC	110	
886	DRGNGSPGAPGAPGHGPGPV-----	915	Db	886	DRGNGSPGAPGAPGHGPGPV-----	915	
111	DKSTQTSPSPCQA 123		Qy	111	DKSTQTSPSPCQA 123		
916	DRGESEGPAGPAGA 928		Db	916	DRGESEGPAGPAGA 928		
			RESULT	111			
			US-09-247-155-178				
			; Sequence 178, Application US/09247155A				
			; Patent No. 6312932				
			; GENERAL INFORMATION				

APPLICANT: Dumas Milne Edwards , Jean-Baptiste
 APPLICANT: Ducleurt, Aymeric
 APPLICANT: Bougnelere, Lydie
 TITLE OF INVENTION: Complementary DNAs
 FILE REFERENCE: GENSET_021A
 CURRENT APPLICATION NUMBER: US/09/247,155A
 CURRENT FILING DATE: 1999-02-09
 EARLIER APPLICATION NUMBER: 60/074,121
 EARLIER FILING DATE: 1998-02-09
 EARLIER APPLICATION NUMBER: 60/081,563
 EARLIER FILING DATE: 1998-04-13
 EARLIER APPLICATION NUMBER: 60/096,116
 EARLIER FILING DATE: 1998-08-10
 EARLIER APPLICATION NUMBER: 60/099,273
 SOFTWARE: Patent .pm
 SEQ ID NO: 178
 LENGTH: 280
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SIGNAL
 LOCATION: -18 .. -1
 US-09-247,155-178

Query Match 8.9%; Score 94.5%; DB 4; Length 280;
 Best Local Similarity 25.0%; Pred. No. 0.056;
 Matches 40; Conservative 16; Mismatches 61; Indels 43; Gaps 6;

Qy 1 MAKOPSDYSSCDEGRQLQPAERPPQLRPGAPTSIQLTEPOQ-----NPEGNHGEG 52
 Db 25 LCRAPDGKGEAAGRGRGRGLKQESEGP-GIRIGIQLKGKDQEPGP-GNPGRVG 83
 Qy 53 DSCPHGSPQPLAPPASPOPFPATR-----PLFIMRSSLIS----- 90
 Db 84 -YPGPSGLGARGIPKIKTGSPGNIKQDQPRPAASARRNNPPIGGNVYFDTVITN 139
 Qy 91 -RSSGYFSTD-----TDRSPAMSCDKSTQPSPPQ 122
 Db 140 QEEPYQNHSGRFVCTVPATTSPSRCCPSGKSACPSSPPQ 179

RESULT 12
 US-09-041-886-3
 Sequence 23, Application US/09041886
 Patent No. 6235872
 GENERAL INFORMATION:
 APPLICANT: Bredesen, Dale E.
 APPLICANT: Rabizadeh, Sharroz
 TITLE OF INVENTION: Proapoptotic Peptides, Dependence
 NUMBER OF SEQUENCES: 72
 ADDRESSEE: Campbell & Flores LLP
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: United States
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/041,886
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LJ 2626
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-9949
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1185 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-041-886-23

Query Match 8.9%; Score 94.5%; DB 4; Length 1185;
 Best Local Similarity 31.0%; Pred. No. 0.36;
 Matches 35; Conservative 8; Mismatches 37; Indels 33; Gaps 7;

Qy 25 PPQLRPGAPTSIQLTEPOQNPPEGNHG-----GEGDSCPHGSP----- 67
 Db 649 PPGYKPGSPPSFRN-----GTPGTYRGTSPAGPGTFKP-GSPTYGPGPPAGPSGLPSP 704

Qy 68 ASPCPFATRSPFLIFMRRSSLLSRSSSYFSFTDRSPAPMSCDKSTQTPSP 120
 Db 705 PPPAAPASGPPL-----SATQIKQEPAAEYETPESPPV-----PARSPSPPP 745

RESULT 13
 US-08-906-865-1
 Sequence 1, Application US/08906865
 Patent No. 6040168
 GENERAL INFORMATION:
 APPLICANT: Greenbard, Paul
 APPLICANT: Burton, Barbara
 APPLICANT: Kao, Hung-Teh
 TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David A. Jackson, Esq.
 STREET: 411 Hackensack Ave, Continental Plaza, 4th
 STREET: Floor
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/906,865
 FILING DATE:
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-202
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-487-3800
 TELEFAX: 201-343-1684
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 580 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 DESCRIPTION: /desc = "Synapsin III"
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-08-906-865-1

Query Match 8.8%; Score 94; DB 3; Length 580;
 Best Local Similarity 27.5%; Pred. No. 0.16;
 Matches 52; Conservative 24; Mismatches 61; Indels 52; Gaps 13;

Qy 2 AKOPSDVSECDREGROL-QPAERPPQLRGAFTSLOT-EPOQ-GNPEGNHGEGDSCP 57
 Db 417 AKSPGQA---OLGPQXQPKRPPP-QGGRQAOQPQPKRPP-SQQR 460

Qy 58 GSPQG--PLAPPASPQPFATRSPLFIMMRSSLLSRSSGGYFSFDTDR----SPAM 108
 Db 461 LSPQGQPLS-PQSGSPQQRSP----GSPQLSASSGSPNQASKPGATLASQPRP 513

Qy 109 SCDKST-----QTPSPP----CQAFNHYLUSAMASMRQAFPADMRPENIAQELRRI 155
 Db 514 VGRSTSQQGEESKKPAPPHUNKSQSLTSQSLTSQRTSPSEDEAK--AEITRNL 570

Qy 156 GDEFNAYYA 164
 Db 571 RKSEASLPS 579

RESULT 14
 Sequence 6, Application US/08466589
 ; Patent No. 5837489

GENERAL INFORMATION:
 ; APPLICANT: Elliot, Kathryn J.
 ; APPLICANT: Ellis, Steven B.
 ; APPLICANT: Harpold, Michael M.
 ; TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
 ; TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Brown, Martin, Haller & McClain
 ; STREET: 1660 Union Street
 ; CITY: San Diego
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92101-2926
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/466,589
 ; FILING DATE: June 5, 1995
 ; CLASSIFICATION: 536
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/028,031
 ; FILING DATE: March 8, 1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie L.
 ; REGISTRATION NUMBER: 33,779
 ; REFERENCE/DOCKET NUMBER: 63362-9950
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-238-0999
 ; TELEFAX: 619-238-0052
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 627 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-466-589-6

Matches 32; Conservative 14; Mismatches 45; Indels 34; Gaps 5;
 Qy 1 MAKOPSDVSECDREGROL-QPAERPPQLRGAFTSLOT-EPOQ-PEGNHGEGDSCP 60
 Db 357 LMKRPSVVKDNCRLLIESHKMASAPRW-----EPEGEPPATSGTQSLHPPSPF 408

Qy 61 QGPLAPASPQPFATRSPLFIMMRSSLLSRSSGGYFSFDTDRS-APMSCDKSTQTPS 118
 Db 409 CVPLDVAEPGP-SCKSP-----SDQLPPQPKPLEAKRDSHPS 445

Query Match 8.8%; Score 94; DB 2; Length 627;
 Best Local Similarity 25.6%; Pred. No. 0.18;
 Matches 32; Conservative 14; Mismatches 45; Indels 34; Gaps 5;
 Qy 1 MAKOPSDVSECDREGROL-QPAERPPQLRGAFTSLOT-EPOQ-PEGNHGEGDSCP 60
 Db 357 LMKRPSVVKDNCRLLIESHKMASAPRW-----EPEGEPPATSGTQSLHPPSPF 408

Qy 61 QGPLAPASPQPFATRSPLFIMMRSSLLSRSSGGYFSFDTDRS-APMSCDKSTQTPS 118
 Db 409 CVPLDVAEPGP-SCKSP-----SDQLPPQPKPLEAKRDSHPS 445

Qy 119 P-PCQ 122
| 11:
Db 446 FGPCR 450

Search completed: August 16, 2002, 10:01:20
Job time: 238 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Result No.	Score	Query	Match	Length	DB	ID	Description
1	111.5	10.5	886	2	A54442		3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17) 4A, cAMP-specific, long splice N: Contains: 3', 5'-cyclic AMP phosphodiesterase HPD4A6 splice form
2	110.5	10.4	472	2	T0169		C; Species: Homo sapiens (man)
3	105.5	9.9	574	2	T4556		C; Accession: A54442; S55788; A36317; S55787
4	105.5	9.9	574	2	T3819		R; Bolger, G.; Michaeli, T.; Martins, T.; St. John, T.; Steiner, B.; Rodgers, L.; Rigg, Mol. Cell. Biol. 13, 6558-6571, 1993
5	103.5	9.7	556	2	T42100		A; Title: A family of human phosphodiesterases homologous to the duncle learning and me
6	103.5	9.7	556	2	T36502		A; Reference number: A54442; MUID:94019330
7	102.5	9.6	1332	1	T48314		A; Accession: A54442
8	100	9.4	99	1	CGREHS		A; Status: translated from GB/EMBL/DBJ
9	100	9.4	627	2	JCA021		A; Molecule type: mRNA
10	99	9.3	636	2	S41067		A; Residues: 1-886 <RES>
11	99	9.3	1464	1	S59856		A; Cross-references: GB:L20965; NID:9347119; PIDN:AAA03588.1; PID:9347120
12	99	9.3	1466	1	CGHUTL		R; Sullivan, M.; Egerton, M.; Shakur, Y.; Marquardsen, A.; Houslay, M.D., Cell. Signal. 6, 793-812, 1994
13	97.5	9.2	289	2	T26812		A; Reference number: S55788; MUID:95194817
14	97.5	9.2	443	2	I49140		A; Accession: S55788
15	97	9.1	852	2	T10811		A; Status: preliminary
16	96	9.0	870	2	G01974		A; Molecule type: mRNA
17	96	9.0	1315	2	B56101		A; Residues: 1-516, 'Y', 518-722, 'R', 724-726, 'R', 728-735, 'E', 737-788, 'E', 7
18	96	9.0	1774	2	B56101		A; Cross-references: GB:M37744
19	95.5	9.0	172	2	D41132		C; Genetics:
20	95.5	9.0	210	2	B44984		A; Gene: GDB:PDE4A; DPDE2
21	95	8.9	298	2	JC1448		A; Cross-references: GDB:138776; OMIM:600126
22	95	8.9	844	2	I58865		A; Map Position: 19p13.1-19q12
23	94.5	8.9	573	2	B70726		C; Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase homology
24	94.5	8.9	628	2	J00110		C; Keywords: alternative splicing; cAMP binding; phosphoric diester hydrolase
25	94.5	8.9	684	2	A53019		F; 432-660/Domain: 3', 5'-cyclic-nucleotide phosphodiesterase homology <CNP>
26	94.5	8.9	1184	2	G01763		A; Query Match: 10.5%; Score: 111.5; DB: 2; Length: 886;
27	94.5	8.9	1184	2	S50832		B; Best Local Similarity: 27.5%; Pred. No. 0.27; Mismatches: 18; Indels: 39; Gaps: 7;
28	94.5	8.9	1234	2	A34911		C; Probable SED - My hypothetical 69K p
29	94	8.8	268	2	H84684		collagen alpha 1(X) p

R; Mamalaki, A.; Remoundos, M.; Tzartos, S.
 submitted to the EMBL Data Library, May 1995
 A; Description: Molecular cloning of human neuronal nicotinic acetylcholine receptor 4-lin
 A; Reference number: S55471
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Gene: GDB:CHRNA4
 A; Cross-references: EMBL:X87629; NID:q854158; PIDN:CAA60959.1; PID:9854159
 A; Residues: 26-627 <NAM>
 A; Map position: 20q13.2-20q13.3
 C; Superfamily: acetylcholine receptor
 C; Keywords: ion channel; neurotransmitter receptor; postsynaptic membrane; transmembrane

Query Match 9.4%; Score 100; DB 2; Length 627;
 Best Local Similarity 26.4%; Pred. No. 1.5;
 Matches 33; Conservative 15; Mismatches 43; Indels 34; Gaps 5;

Qy 1 MAKOPSDVSSCDREGRQLOPAERPPQLRGAPTSLOTEPQGNPREGNNGEGDSCPAGGSP 60
 Db 357 LMKRPVSKVNDCRLIESMKMASPRFWP-----EPFEGEPATPSLPPSPS 408

Qy 61 QGPLAPPASPQPFATRSPSPLEFMRSSSLSRSSGYFSSFTDRSP--AMMSCDKSTQTPS 118
 Db 409 CVPLDVPAEPQP-SCRSKPS-----SDQLPPQOLEAERASPHS 445

Qy 119 P-PCQ 122
 Db 446 PGPCR 450

RESULT 10
 S41067
 collagen alpha 1(III) chain - rat
 C; Species: Rattus norvegicus (Norway rat)
 C; Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
 C; Accession: S41067
 R; Glumoff, V.; Mäkelä, J. K.; Vuorio, E.
 A; Title: Cloning of cDNA for rat pro alpha-1(III) collagen mRNA. Different expression patterns of collagen alpha-1(III) and alpha-2(IV) collagen mRNAs by esterase
 A; Reference number: S41067; MUID:94114571
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 1-636 <GLU>
 A; Cross-references: EMBL:X70369; NID:957915; PIDN:CAA49832.1; PID:957916
 R; Frankel, F.R.; Hsu, C.Y.J.; Myers, J.C.; Lin, E.; Lytle, C.R.; Komm, B.; Mohn, K.
 DNA 7, 347-354, 1988
 A; Title: Cloning of alpha-2 (I), alpha-1 (III), and alpha-2(IV) collagen mRNAs by esterase
 A; Reference number: A29905; MUID:88296083
 A; Accession: S41067
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 308-482 <FRA>
 A; Cross-references: EMBL:M21354; NID:9203500; PIDN:AAA40942.1; PID:9203501
 R; Glumoff, V.; Mäkelä, J. K.; Vuorio, E.
 submitted to the EMBL Data Library, February 1993
 A; Reference number: S31924
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 2-636 <GL2>
 A; Cross-references: EMBL:X70369
 C; Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 F; 408-636/domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 9.3%; Score 99; DB 2; Length 636;
 Best Local Similarity 35.4%; Pred. No. 1.9;
 Matches 29; Conservative 8; Mismatches 31; Indels 14; Gaps

submitted to the EMBL Data Library, February 1989

A: Reference number: S59511; MUID: 96067614

A; Accession: S59511
 A; Molecule type: mRNA
 A; Residues: 302-423 <CHI>
 A; Cross-references: GB:S79877; NID:q1195576; PIDN:AAB35615.1; PID:q1195577
 R; Seyer, J.M.; Kang, A.H.
 Biochemistry 17, 3404-3411, 1978
 A; Title: Covalent structure of collagen: amino acid sequence of five consecutive
 A; Reference number: A90414; MUID:79000343
 A; Accession: A90414
 A; Molecule type: protein
 A; Residues: 399-675, 'N', 677-727 <SEY3>
 A; Experimental source: liver
 R; Lee, B.; Vitale, E.; Supersti-Furga, A.; Steinmann, B.; Ramirez, F.
 J. Biol. Chem. 266, 5236-5239, 1991
 A; Title: G to T transversion at position +5 of a splice donor site causes skip
 A; Reference number: I55349; MUID:91161621
 A; Accession: I55349
 A; Status: translated from GB/EMBL/DBJ
 A; Molecule type: DNA
 A; Residues: 537-605 <LEB>
 A; Cross-references: GB:M59312; NID:q180815; PIDN:AAA52041.1; PID:q180816
 R; Seyer, J.M.; Mainardi, C.; Kang, A.H.
 Biochemistry 19, 1583-1589, 1980
 A; Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-C
 A; Reference number: A90438; MUID:80198282
 A; Accession: A90438
 A; Molecule type: protein
 A; Residues: 728-895, 'A', 897-964 <SEY4>
 A; Experimental source: liver
 R; Cole, W.G.; Chiodo, A.A.; Lamande, S.R.; Janecekko, R.; Ramirez, F.; Dahl, H.
 J. Biol. Chem. 265, 17070-17077, 1990
 A; Title: A base substitution at a splice site in the COL3A1 gene causes exon 8
 A; Reference number: A38303; MUID:91009133
 A; Accession: A38303
 A; Molecule type: mRNA
 A; Residues: 861-1015 <COL>
 A; Cross-references: GB:J05617; GB:MM5603; GB:MS9227; NID:q180878; PIDN:AAB5938
 A; Note: a mutant sequence: 942-977 spiced out from a patient with Ehlers-
 Danlos syndrome
 R; Mankoo, B.S.; Dagleish, R.
 Nucleic Acids Res. 16, 2337, 1988
 A; Title: Human pro alpha1 (III) collagen: cDNA sequence for the 3' end.
 A; Reference number: S02119; MUID:8818827
 A; Accession: S02119
 A; Status: translation not shown
 A; Molecule type: mRNA
 A; Residues: 950-1018, 'Y', 1020-1183, 'S', 1185-1466 <MAN>
 A; Cross-references: EMBL:X06700; NID:930053; PIDN:CAA29886.1; PID:930054
 R; Seyer, J.M.; Kang, A.H.
 Biochemistry 20, 2621-2627, 1981
 A; Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-C
 A; Reference number: A90446; MUID:81200139
 A; Accession: A90446
 A; Molecule type: protein
 A; Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155
 A; Experimental source: liver
 R; Lordi, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Mortow, S.; Rosenblow,
 R; Miskulin, M.; Dagleish, R.; Klue-Beckerman, B.; Rennard, S.I.; Tolstoshev,
 Biochemistry 25, 1408-1413, 1986
 A; Title: Human type III collagen gene expression is coordinately modulated with
 A; Reference number: I55393; MUID:85087944
 A; Accession: A93551
 A; Molecule type: mRNA
 A; Residues: 1065-1155, 'P', 1157-1466 <LOI>
 A; Cross-references: EMBL:X01655; EMBL:X01742; NID:q29584; PIDN:CAA25821.1
 R; Seyer, J.M.; Kang, A.H.
 Biochemistry 12, 9383-9394, 1984
 A; Title: Molecular cloning and carboxyl-propeptide analysis of human type III
 A; Reference number: I55393; MUID:86187804
 A; Accession: A93551
 A; Molecule type: mRNA
 A; Residues: 1161-1200 <MIS>
 A; Cross-references: GB:MM1346; NID:q180415; PIDN:AAA52003.1; PID:q180416

C;Accession: T10811	R;Irie, M.; Hata, Y.; Takai, Y.	A;Status: preliminary; translated from GB/EMBL/DDJB
A;Accession: T10811	A;Status: preliminary; translated from GB/EMBL/DDJB	A;Description: Cloning of new isoforms of PSD-95/SAP90 related genes.
A;Accession: T10811	A;Status: preliminary; translated from GB/EMBL/DDJB	A;Reference number: Z17166
A;Accession: T10811	A;Status: preliminary; translated from GB/EMBL/DDJB	A;Cross-references: 1-852 <URI>
A;Accession: T10811	A;Status: preliminary; translated from GB/EMBL/DDJB	C;Genetics:
A;Accession: T10811	A;Status: preliminary; translated from GB/EMBL/DDJB	C;Note: PSD-95/SAP90 related gene 2 (chapsyn isoform2)
A;Accession: T10811	A;Status: preliminary; translated from GB/EMBL/DDJB	C;Superfamily: discs-large tumor suppressor; GLGF domain homology; guanylate kinase homology
A;Accession: T10811	A;Status: preliminary; translated from GB/EMBL/DDJB	F:198-276/Domain: GLGF domain homology <GLG>
A;Accession: T10811	A;Status: preliminary; translated from GB/EMBL/DDJB	F:543-601/Domain: SH3 homology <SH3>
A;Accession: T10811	A;Status: preliminary; translated from GB/EMBL/DDJB	F:663-840/Domain: guanylate kinase homology <GK1>
Query Match	9.1%	Score 97; DB 2; Length 852;
Best Local Similarity	23.8%	Pred. No. 3.7;
Matches 44;	Conservative 21;	Mismatches 58; Indels 62; Gaps 9;
QY	28	LRPGAPTSL-QTEPGGNPEGAHGGEDSCPH -GSEQGPL-----APPASCPGPFA---- 74
Db	274	LKVGRPTTIVMDP1GPPDTHSYSPPMENHLNSGNGNTLEYKTSLLPISGPRYSP1PKH 333
QY	75	-----TSPLEFIMRQSSLLSRSSSGYFSFDTDRSPAMSCDKS -TQTPSPPPCQAFN 125
Db	334	MLVEDDEYTRPEPVSTVNLCDKPDS-----PRHYSPEVCDKSFLLSTYR----- 380
QY	126	HYL-----SAMASMRQAEADMRRPBIWIQELRRIGDEFNAYYARRVFLNNYQAAEDHP 179
Db	381	HYHGLLPPDSMDTSHSQHSTATROPSVTLQRAI-----SLEGEPE 419
QY	180	RMVIL 184
QY	420	----- 420

Search completed: August 16, 2002, 10:01:47

THIS PAGE BLANK (USPTO)

Scoring table:	BLOSUM62	Gapop 10.0 , Gapext 0.5	
Searched:	105224 seqs., 38719550 residues		
total number of hits satisfying chosen parameters:	105224		
Minimum DB seq length:	0		
Maximum DB seq length:	20000000000		
Post-processing:	Minimum Match 0%		
	Maximum Match 100%		
Database :	Swissprot_40::*	Listing first 45 summaries	
Pred. No.	Score	Query Match	Description
No.		Length	DB ID
1	1065	100.0	1 BIM_HUMAN
2	908	85.3	1 BIM_MOUSE
3	893	83.8	1 BIM_RAT
4	111	10.5	1 CN4A_HUMAN
5	106.5	10.0	1 EXON_HSV2
6	102.5	9.6	1 CUT1_MOUSE
7	100	9.4	1 ACH4_HUMAN
8	99	9.3	1 CAL1_RAT
9	99	9.3	1 CAL1_MOUSE
10	99	9.3	1 CAL13_HUMAN
11	98	9.2	1 COL1_HSV57
12	97.5	9.2	1 CUT1_HUMAN
13	97	9.1	1 DLG2_RAT
14	96	9.0	1 DLG2_HUMAN
15	96	9.0	1 CA1H_MOUSE
16	95	9.0	1 CAC2_HAECKO
17	95	8.9	1 CC34_CAEEL
18	95	8.9	1 CN4A_RAT
19	95	8.9	1516 1 CA1H_HUMAN
20	94.5	8.9	573 1 SECD_MYTU
21	94.5	8.9	628 1 V7OK_TYAMA
22	94.5	8.9	1183 1 DRPL_RAT
23	94.5	8.9	1185 1 DRPL_HUMAN
24	94	8.9	1234 1 B3A2_RAT
25	94	8.8	580 1 SYN3_HUMAN
26	94	8.8	1460 1 CA11_CANFA
27	93.5	8.8	902 1 NFCA_HUMAN
28	92.5	8.7	245 1 C1Q4_HUMAN
29	92.5	8.7	408 1 AL_DROME
30	92.5	8.7	1446 1 IE18_PRVTF
31	92.5	8.7	1461 1 IE18_PRVTF
32	91.5	8.6	234 1 PRPM_HUMAN
33	91.5	8.6	276 1 PRPL_HUMAN

RN	[2]	SEQUENCE OF 64-1395 FROM N.A. (ISOFORM 1).	Score 9.68; Pred. No. 3.5;
RP	RC	STRAINA/J. AND BALBC; TISSUE:Brain;	Length 1395;
RC	RC	MEDLINE=94244481; PubMed=9710552;	Matches 54; Conservative 25; Mismatches 76; Indels 73; Gaps 9;
RX	RC	Varclerc I., Tissier-Seta J.-P., Hirsch M.R., Martinez S., Goridis C., Brunet J.-F.;	
RA	RA	"The mouse homeodomain protein Phox2 regulates Ncam promoter activity in concert with Cux/Cbp and is a putative determinant of neurotransmitter phenotype.";	
RT	RT	Development 119:881-896(1993).	
RN	[3]	SEQUENCE OF 642-1395 FROM N.A.	
RP	RP	MEDLINE=96437626; PubMed=8840273;	
RA	RA	Den Heuvel G.B., Bodmer R., McConnel K.R., Nagami G.T., Igarashi P.; "Expression of a cut-related homeobox gene in developing and polycystic mouse kidney".	
RT	RT	Kidney Int. 50:453-461(1996).	
RN	[4]	SEQUENCE OF 936-1395 FROM N.A.	
RP	RP	Quaggan S.E., Igarashi P.;	
RA	RA	"A unique variant of a homeobox gene related to Drosophila cut is expressed in mouse testis"; Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.	
RL	RL	-1- FUNCTION: PROBABLY HAS A BROAD ROLE IN MAMMALIAN DEVELOPMENT AS A REPRESSOR OF DEVELOPMENTALLY REGULATED GENE EXPRESSION. MAY ACT BY PREVENTING BLINDING OF POSITIVELY-ACTIVATING CCAAT FACTORS TO PROMOTERS (BY SIMILARITY). COMPONENT OF NF-MUNR REPRESSOR: BINDS TO THE MARS (5, AND 3) OF THE IMMUNOGLOBULIN HEAVY CHAIN ENHANCER.	
CC	CC	-1- SUBCELLULAR LOCATION: Nuclear.	
CC	CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE PRODUCED BY ALTERNATIVE SPLICING.	
CC	CC	-1- SIMILARITY: CONTAINS 3 CUT DOMAINS.	
CC	CC	-1- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEBOX PROTEINS.	
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
DR	DR	EMBL: AF004225; AAD12485.1; -.	
DR	DR	EMBL: X75013; CAA52922.1; -.	
DR	DR	EMBL: U46683; AAC52775.1; -.	
DR	DR	EMBL: U46684; AB41146.1; -.	
DR	DR	HSSP: P10037; IAU7.	
DR	DR	MGD: MGI:888568; Cut11.	
DR	DR	InterPro: IPR0033350; CUT.	
DR	DR	InterPro: IPR0033356; Homeobox.	
DR	DR	Pfam: PF002376; CUT. 3.	
DR	DR	Pfam: PF00046; homeobox. 2.	
DR	DR	SMART: SM00389; HOX. 1.	
DR	DR	PROSITE: PS00027; HOMEBOX.1; 1.	
DR	DR	PROSITE: PS03071; HOMEBOX.2; 1.	
KW	KW	Transcription regulation; Homeobox; DNA-binding; Developmental protein; Nuclear protein; Repeat; Repressor; Coiled coil; Alternative splicing.	
FT	FT	NON-TER 1 1	
FT	FT	DOMAIN 1 243	
FT	FT	DNA-BIND 420 507	
FT	FT	DOMAIN 547 603	
FT	FT	DNA-BIND 809 896	
FT	FT	DNA-BIND 992 1079	
FT	FT	DNA-BIND 1119 1178	
FT	FT	VARSPLIC 287 388	
FT	FT	MISSING (IN ISOFORM 2).	
FT	FT	CONFLICT 1360 1360	
FT	FT	G > A (IN REF. 2).	
FT	FT	CONFLICT 1365 1365	
FT	FT	P > L (IN REF. 1).	
SEQUENCE	RN	AA; 151802 MW; D062CC22/D7A163E CRC64;	
SEQUENCE	RN	1395 AA;	
	RP	SEQUENCE FROM N.A.	

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggaley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blaizey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrlin W., Butler A.P., Carter C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhami P., Dunn M., Ellington A.G., Franklin J.A., Fraser A., French L., Garner P., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Hunt S.E., Huckle E., Hunt A.R., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Lehrvahlaih M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConnaughey L.J., McMurray A.A., Milne S., Misery D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.T., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C.A., Steward C.A., Sulston J.E., Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., Wilmng L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.; RT "The DNA sequence and comparative analysis of human chromosome 20."; RL Nature 414: 865-871(2001).
 RN [6]
 RP SEQUENCE OF 26-627 FROM N.A.
 RC TISSUE=Brain;
 RA Mama-laki A., Remoundos M., Tzartos S.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
 RN [7]
 RX MEDLINE=26024658; PubMed=7550350;
 RA Steinlein O.K., Mulley J.C., Propping P., Wallace R.H., Phillips H.A., Scheffer I.E., Berkovic S.F.;
 RT "A missense mutation in the neuronal nicotinic acetylcholine receptor alpha-4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy.";
 RT Nat. Genet. 11:201-203(1995).
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA MEMBRANE.
 CC -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPE OF SUBUNITS: ALPHA AND NON-ALPHA (BETA).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DISEASE: Defects in CHRNA4 are the cause of autosomal dominant nocturnal frontal lobe epilepsy type 1 (ENFL1 or ADNFE). It is a disease characterized by clusters of motor seizures during sleep.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: L35921; AAA64743.1; DR EMBL: X89741; CAA61893.1; DR EMBL: X89742; CAA61893.1; JOINED.
 CC EMBL: X89743; CAA61893.1; JOINED.
 CC EMBL: X89744; CAA61893.1; JOINED.
 CC EMBL: X89745; CAA61893.1; JOINED.
 CC EMBL: X89746; CAA61893.1; JOINED.
 CC EMBL: X89747; CAA61893.1; JOINED.
 CC EMBL: X89748; CAA61893.1; JOINED.
 CC EMBL: X89749; CAA61893.1; JOINED.
 CC EMBL: X89750; CAA61893.1; JOINED.
 CC EMBL: X89751; CAA61893.1; JOINED.
 CC EMBL: X89752; CAA61893.1; JOINED.
 CC EMBL: X89753; CAA61893.1; JOINED.
 CC EMBL: X89754; CAA61893.1; JOINED.
 CC EMBL: X89755; CAA61893.1; JOINED.
 CC EMBL: X89756; CAA61893.1; JOINED.
 CC EMBL: X89757; CAA61893.1; JOINED.
 CC EMBL: X89758; CAA61893.1; JOINED.
 CC EMBL: X89759; CAA61893.1; JOINED.
 CC EMBL: X89760; CAA61893.1; JOINED.
 CC EMBL: X89761; CAA61893.1; JOINED.
 CC EMBL: X89762; CAA61893.1; JOINED.
 CC EMBL: X89763; CAA61893.1; JOINED.
 CC EMBL: X89764; CAA61893.1; JOINED.
 CC EMBL: X89765; CAA61893.1; JOINED.
 CC EMBL: X89766; CAA61893.1; JOINED.
 CC EMBL: X89767; CAA61893.1; JOINED.
 CC EMBL: X89768; CAA61893.1; JOINED.
 CC EMBL: X89769; CAA61893.1; JOINED.
 CC EMBL: X89770; CAA61893.1; JOINED.
 CC EMBL: X89771; CAA61893.1; JOINED.
 CC EMBL: X89772; CAA61893.1; JOINED.
 CC EMBL: X89773; CAA61893.1; JOINED.
 CC EMBL: X89774; CAA61893.1; JOINED.
 CC EMBL: X89775; CAA61893.1; JOINED.
 CC EMBL: X89776; CAA61893.1; JOINED.
 CC EMBL: X89777; CAA61893.1; JOINED.
 CC EMBL: X89778; CAA61893.1; JOINED.
 CC EMBL: X89779; CAA61893.1; JOINED.
 CC EMBL: X89780; CAA61893.1; JOINED.
 CC EMBL: X89781; CAA61893.1; JOINED.
 CC EMBL: X89782; CAA61893.1; JOINED.
 CC EMBL: X89783; CAA61893.1; JOINED.
 CC EMBL: X89784; CAA61893.1; JOINED.
 CC EMBL: X89785; CAA61893.1; JOINED.
 CC EMBL: X89786; CAA61893.1; JOINED.
 CC EMBL: X89787; CAA61893.1; JOINED.
 CC EMBL: X89788; CAA61893.1; JOINED.
 CC EMBL: X89789; CAA61893.1; JOINED.
 CC EMBL: X89790; CAA61893.1; JOINED.
 CC EMBL: X89791; CAA61893.1; JOINED.
 CC EMBL: X89792; CAA61893.1; JOINED.
 CC EMBL: X89793; CAA61893.1; JOINED.
 CC EMBL: X89794; CAA61893.1; JOINED.
 CC EMBL: X89795; CAA61893.1; JOINED.
 CC EMBL: X89796; CAA61893.1; JOINED.
 CC EMBL: X89797; CAA61893.1; JOINED.
 CC EMBL: X89798; CAA61893.1; JOINED.
 CC EMBL: X89799; CAA61893.1; JOINED.
 CC EMBL: X89800; CAA61893.1; JOINED.
 CC EMBL: X89801; CAA61893.1; JOINED.
 CC EMBL: X89802; CAA61893.1; JOINED.
 CC EMBL: X89803; CAA61893.1; JOINED.
 CC EMBL: X89804; CAA61893.1; JOINED.
 CC EMBL: X89805; CAA61893.1; JOINED.
 CC EMBL: X89806; CAA61893.1; JOINED.
 CC EMBL: X89807; CAA61893.1; JOINED.
 CC EMBL: X89808; CAA61893.1; JOINED.
 CC EMBL: X89809; CAA61893.1; JOINED.
 CC EMBL: X89810; CAA61893.1; JOINED.
 CC EMBL: X89811; CAA61893.1; JOINED.
 CC EMBL: X89812; CAA61893.1; JOINED.
 CC EMBL: X89813; CAA61893.1; JOINED.
 CC EMBL: X89814; CAA61893.1; JOINED.
 CC EMBL: X89815; CAA61893.1; JOINED.
 CC EMBL: X89816; CAA61893.1; JOINED.
 CC EMBL: X89817; CAA61893.1; JOINED.
 CC EMBL: X89818; CAA61893.1; JOINED.
 CC EMBL: X89819; CAA61893.1; JOINED.
 CC EMBL: X89820; CAA61893.1; JOINED.
 CC EMBL: X89821; CAA61893.1; JOINED.
 CC EMBL: X89822; CAA61893.1; JOINED.
 CC EMBL: X89823; CAA61893.1; JOINED.
 CC EMBL: X89824; CAA61893.1; JOINED.
 CC EMBL: X89825; CAA61893.1; JOINED.
 CC EMBL: X89826; CAA61893.1; JOINED.
 CC EMBL: X89827; CAA61893.1; JOINED.
 CC EMBL: X89828; CAA61893.1; JOINED.
 CC EMBL: X89829; CAA61893.1; JOINED.
 CC EMBL: X89830; CAA61893.1; JOINED.
 CC EMBL: X89831; CAA61893.1; JOINED.
 CC EMBL: X89832; CAA61893.1; JOINED.
 CC EMBL: X89833; CAA61893.1; JOINED.
 CC EMBL: X89834; CAA61893.1; JOINED.
 CC EMBL: X89835; CAA61893.1; JOINED.
 CC EMBL: X89836; CAA61893.1; JOINED.
 CC EMBL: X89837; CAA61893.1; JOINED.
 CC EMBL: X89838; CAA61893.1; JOINED.
 CC EMBL: X89839; CAA61893.1; JOINED.
 CC EMBL: X89840; CAA61893.1; JOINED.
 CC EMBL: X89841; CAA61893.1; JOINED.
 CC EMBL: X89842; CAA61893.1; JOINED.
 CC EMBL: X89843; CAA61893.1; JOINED.
 CC EMBL: X89844; CAA61893.1; JOINED.
 CC EMBL: X89845; CAA61893.1; JOINED.
 CC EMBL: X89846; CAA61893.1; JOINED.
 CC EMBL: X89847; CAA61893.1; JOINED.
 CC EMBL: X89848; CAA61893.1; JOINED.
 CC EMBL: X89849; CAA61893.1; JOINED.
 CC EMBL: X89850; CAA61893.1; JOINED.
 CC EMBL: X89851; CAA61893.1; JOINED.
 CC EMBL: X89852; CAA61893.1; JOINED.
 CC EMBL: X89853; CAA61893.1; JOINED.
 CC EMBL: X89854; CAA61893.1; JOINED.
 CC EMBL: X89855; CAA61893.1; JOINED.
 CC EMBL: X89856; CAA61893.1; JOINED.
 CC EMBL: X89857; CAA61893.1; JOINED.
 CC EMBL: X89858; CAA61893.1; JOINED.
 CC EMBL: X89859; CAA61893.1; JOINED.
 CC EMBL: X89860; CAA61893.1; JOINED.
 CC EMBL: X89861; CAA61893.1; JOINED.
 CC EMBL: X89862; CAA61893.1; JOINED.
 CC EMBL: X89863; CAA61893.1; JOINED.
 CC EMBL: X89864; CAA61893.1; JOINED.
 CC EMBL: X89865; CAA61893.1; JOINED.
 CC EMBL: X89866; CAA61893.1; JOINED.
 CC EMBL: X89867; CAA61893.1; JOINED.
 CC EMBL: X89868; CAA61893.1; JOINED.
 CC EMBL: X89869; CAA61893.1; JOINED.
 CC EMBL: X89870; CAA61893.1; JOINED.
 CC EMBL: X89871; CAA61893.1; JOINED.
 CC EMBL: X89872; CAA61893.1; JOINED.
 CC EMBL: X89873; CAA61893.1; JOINED.
 CC EMBL: X89874; CAA61893.1; JOINED.
 CC EMBL: X89875; CAA61893.1; JOINED.
 CC EMBL: X89876; CAA61893.1; JOINED.
 CC EMBL: X89877; CAA61893.1; JOINED.
 CC EMBL: X89878; CAA61893.1; JOINED.
 CC EMBL: X89879; CAA61893.1; JOINED.
 CC EMBL: X89880; CAA61893.1; JOINED.
 CC EMBL: X89881; CAA61893.1; JOINED.
 CC EMBL: X89882; CAA61893.1; JOINED.
 CC EMBL: X89883; CAA61893.1; JOINED.
 CC EMBL: X89884; CAA61893.1; JOINED.
 CC EMBL: X89885; CAA61893.1; JOINED.
 CC EMBL: X89886; CAA61893.1; JOINED.
 CC EMBL: X89887; CAA61893.1; JOINED.
 CC EMBL: X89888; CAA61893.1; JOINED.
 CC EMBL: X89889; CAA61893.1; JOINED.
 CC EMBL: X89890; CAA61893.1; JOINED.
 CC EMBL: X89891; CAA61893.1; JOINED.
 CC EMBL: X89892; CAA61893.1; JOINED.
 CC EMBL: X89893; CAA61893.1; JOINED.
 CC EMBL: X89894; CAA61893.1; JOINED.
 CC EMBL: X89895; CAA61893.1; JOINED.
 CC EMBL: X89896; CAA61893.1; JOINED.
 CC EMBL: X89897; CAA61893.1; JOINED.
 CC EMBL: X89898; CAA61893.1; JOINED.
 CC EMBL: X89899; CAA61893.1; JOINED.
 CC EMBL: X89900; CAA61893.1; JOINED.
 CC EMBL: X89901; CAA61893.1; JOINED.
 CC EMBL: X89902; CAA61893.1; JOINED.
 CC EMBL: X89903; CAA61893.1; JOINED.
 CC EMBL: X89904; CAA61893.1; JOINED.
 CC EMBL: X89905; CAA61893.1; JOINED.
 CC EMBL: X89906; CAA61893.1; JOINED.
 CC EMBL: X89907; CAA61893.1; JOINED.
 CC EMBL: X89908; CAA61893.1; JOINED.
 CC EMBL: X89909; CAA61893.1; JOINED.
 CC EMBL: X89910; CAA61893.1; JOINED.
 CC EMBL: X89911; CAA61893.1; JOINED.
 CC EMBL: X89912; CAA61893.1; JOINED.
 CC EMBL: X89913; CAA61893.1; JOINED.
 CC EMBL: X89914; CAA61893.1; JOINED.
 CC EMBL: X89915; CAA61893.1; JOINED.
 CC EMBL: X89916; CAA61893.1; JOINED.
 CC EMBL: X89917; CAA61893.1; JOINED.
 CC EMBL: X89918; CAA61893.1; JOINED.
 CC EMBL: X89919; CAA61893.1; JOINED.
 CC EMBL: X89920; CAA61893.1; JOINED.
 CC EMBL: X89921; CAA61893.1; JOINED.
 CC EMBL: X89922; CAA61893.1; JOINED.
 CC EMBL: X89923; CAA61893.1; JOINED.
 CC EMBL: X89924; CAA61893.1; JOINED.
 CC EMBL: X89925; CAA61893.1; JOINED.
 CC EMBL: X89926; CAA61893.1; JOINED.
 CC EMBL: X89927; CAA61893.1; JOINED.
 CC EMBL: X89928; CAA61893.1; JOINED.
 CC EMBL: X89929; CAA61893.1; JOINED.
 CC EMBL: X89930; CAA61893.1; JOINED.
 CC EMBL: X89931; CAA61893.1; JOINED.
 CC EMBL: X89932; CAA61893.1; JOINED.
 CC EMBL: X89933; CAA61893.1; JOINED.
 CC EMBL: X89934; CAA61893.1; JOINED.
 CC EMBL: X89935; CAA61893.1; JOINED.
 CC EMBL: X89936; CAA61893.1; JOINED.
 CC EMBL: X89937; CAA61893.1; JOINED.
 CC EMBL: X89938; CAA61893.1; JOINED.
 CC EMBL: X89939; CAA61893.1; JOINED.
 CC EMBL: X89940; CAA61893.1; JOINED.
 CC EMBL: X89941; CAA61893.1; JOINED.
 CC EMBL: X89942; CAA61893.1; JOINED.
 CC EMBL: X89943; CAA61893.1; JOINED.
 CC EMBL: X89944; CAA61893.1; JOINED.
 CC EMBL: X89945; CAA61893.1; JOINED.
 CC EMBL: X89946; CAA61893.1; JOINED.
 CC EMBL: X89947; CAA61893.1; JOINED.
 CC EMBL: X89948; CAA61893.1; JOINED.
 CC EMBL: X89949; CAA61893.1; JOINED.
 CC EMBL: X89950; CAA61893.1; JOINED.
 CC EMBL: X89951; CAA61893.1; JOINED.
 CC EMBL: X89952; CAA61893.1; JOINED.
 CC EMBL: X89953; CAA61893.1; JOINED.
 CC EMBL: X89954; CAA61893.1; JOINED.
 CC EMBL: X89955; CAA61893.1; JOINED.
 CC EMBL: X89956; CAA61893.1; JOINED.
 CC EMBL: X89957; CAA61893.1; JOINED.
 CC EMBL: X89958; CAA61893.1; JOINED.
 CC EMBL: X89959; CAA61893.1; JOINED.
 CC EMBL: X89960; CAA61893.1; JOINED.
 CC EMBL: X89961; CAA61893.1; JOINED.
 CC EMBL: X89962; CAA61893.1; JOINED.
 CC EMBL: X89963; CAA61893.1; JOINED.
 CC EMBL: X89964; CAA61893.1; JOINED.
 CC EMBL: X89965; CAA61893.1; JOINED.
 CC EMBL: X89966; CAA61893.1; JOINED.
 CC EMBL: X89967; CAA61893.1; JOINED.
 CC EMBL: X89968; CAA61893.1; JOINED.
 CC EMBL: X89969; CAA61893.1; JOINED.
 CC EMBL: X89970; CAA61893.1; JOINED.
 CC EMBL: X89971; CAA61893.1; JOINED.
 CC EMBL: X89972; CAA61893.1; JOINED.
 CC EMBL: X89973; CAA61893.1; JOINED.
 CC EMBL: X89974; CAA61893.1; JOINED.
 CC EMBL: X89975; CAA61893.1; JOINED.
 CC EMBL: X89976; CAA61893.1; JOINED.
 CC EMBL: X89977; CAA61893.1; JOINED.
 CC EMBL: X89978; CAA61893.1; JOINED.
 CC EMBL: X89979; CAA61893.1; JOINED.
 CC EMBL: X89980; CAA61893.1; JOINED.
 CC EMBL: X89981; CAA61893.1; JOINED.
 CC EMBL: X89982; CAA61893.1; JOINED.
 CC EMBL: X89983; CAA61893.1; JOINED.
 CC EMBL: X89984; CAA61893.1; JOINED.
 CC EMBL: X89985; CAA61893.1; JOINED.
 CC EMBL: X89986; CAA61893.1; JOINED.
 CC EMBL: X89987; CAA61893.1; JOINED.
 CC EMBL: X89988; CAA61893.1; JOINED.
 CC EMBL: X89989; CAA61893.1; JOINED.
 CC EMBL: X89990; CAA61893.1; JOINED.
 CC EMBL: X89991; CAA61893.1; JOINED.
 CC EMBL: X89992; CAA61893.1; JOINED.
 CC EMBL: X89993; CAA61893.1; JOINED.
 CC EMBL: X89994; CAA61893.1; JOINED.
 CC EMBL: X89995; CAA61893.1; JOINED.
 CC EMBL: X89996; CAA61893.1; JOINED.
 CC EMBL: X89997; CAA61893.1; JOINED.
 CC EMBL: X89998; CAA61893.1; JOINED.
 CC EMBL: X89999; CAA61893.1; JOINED.
 CC EMBL: X89900; CAA61893.1; JOINED.
 CC EMBL: X89901; CAA61893.1; JOINED.
 CC EMBL: X89902; CAA61893.1; JOINED.
 CC EMBL: X89903; CAA61893.1; JOINED.
 CC EMBL: X89904; CAA61893.1; JOINED.
 CC EMBL: X89905; CAA61893.1; JOINED.
 CC EMBL: X89906; CAA61893.1; JOINED.
 CC EMBL: X89907; CAA61893.1; JOINED.
 CC EMBL: X89908; CAA61893.1; JOINED.
 CC EMBL: X89909; CAA61893.1; JOINED.
 CC EMBL: X89910; CAA61893.1; JOINED.
 CC EMBL: X89911; CAA61893.1; JOINED.
 CC EMBL: X89912; CAA61893.1; JOINED.
 CC EMBL: X89913; CAA61893.1; JOINED.
 CC EMBL: X89914; CAA61893.1; JOINED.
 CC EMBL: X89915; CAA61893.1; JOINED.
 CC EMBL: X89916; CAA61893.1; JOINED.
 CC EMBL: X89917; CAA61893.1; JOINED.
 CC EMBL: X89918; CAA61893.1; JOINED.
 CC EMBL: X89919; CAA61893.1; JOINED.
 CC EMBL: X89920; CAA61893.1; JOINED.
 CC EMBL: X89921; CAA61893.1; JOINED.
 CC EMBL: X89922; CAA61893.1; JOINED.
 CC EMBL: X89923; CAA61893.1; JOINED.
 CC EMBL: X89924; CAA61893.1; JOINED.
 CC EMBL: X89925; CAA61893.1; JOINED.
 CC EMBL: X89926; CAA61893.1; JOINED.
 CC EMBL: X89927; CAA61893.1; JOINED.
 CC EMBL: X89928; CAA61893.1; JOINED.
 CC EMBL: X89929; CAA61893.1; JOINED.
 CC EMBL: X89930; CAA61893.1; JOINED.
 CC EMBL: X89931; CAA61893.1; JOINED.
 CC EMBL: X89932; CAA61893.1; JOINED.
 CC EMBL: X89933; CAA61893.1; JOINED.
 CC EMBL: X89934; CAA61893.1; JOINED.
 CC EMBL: X89935; CAA61893.1; JOINED.
 CC EMBL: X89936; CAA61893.1; JOINED.
 CC EMBL: X89937; CAA61893.1; JOINED.
 CC EMBL: X89938; CAA61893.1; JOINED.
 CC EMBL: X89939; CAA61893.1; JOINED.
 CC EMBL: X89940; CAA61893.1; JOINED.
 CC EMBL: X89941; CAA61893.1; JOINED.
 CC EMBL: X89942; CAA61893.1; JOINED.
 CC EMBL: X89943; CAA61893.1; JOINED.
 CC EMBL: X89944; CAA61893.1; JOINED.
 CC EMBL: X89945; CAA61893.1; JOINED.
 CC EMBL: X89946; CAA61893.1; JOINED.
 CC EMBL: X89947; CAA61893.1; JOINED.
 CC EMBL: X89948; CAA61893.1; JOINED.
 CC EMBL: X89949; CAA61893.1; JOINED.
 CC EMBL: X89950; CAA61893.1; JOINED.
 CC EMBL: X89951; CAA61893.1; JOINED.
 CC EMBL: X89952; CAA61893.1; JOINED.
 CC EMBL: X89953; CAA61893.1; JOINED.
 CC EMBL: X89954; CAA61893.1; JOINED.
 CC EMBL: X89955; CAA61893.1; JOINED.
 CC EMBL: X89956; CAA61893.1; JOINED.
 CC EMBL: X89957; CAA61893.1; JOINED.
 CC EMBL: X89958; CAA61893.1; JOINED.
 CC EMBL: X89959; CAA61893.1; JOINED.
 CC EMBL: X89960; CAA61893.1; JOINED.
 CC EMBL: X89961; CAA61893.1; JOINED.
 CC EMBL: X89962; CAA61893.1; JOINED.
 CC EMBL: X89963; CAA61893.1; JOINED.
 CC EMBL: X89964; CAA61893.1; JOINED.
 CC EMBL: X89965; CAA61893.1; JOINED.
 CC EMBL: X89966; CAA61893.1; JOINED.
 CC EMBL: X89967; CAA61893.1; JOINED.
 CC EMBL: X89968; CAA61893.1; JOINED.
 CC EMBL: X89969; CAA61893.1; JOINED.
 CC EMBL: X89970; CAA61893.1; JOINED.
 CC EMBL: X89971; CAA61893.1; JOINED.
 CC EMBL: X89972; CAA61893.1; JOINED.
 CC EMBL: X89973; CAA61893.1; JOINED.
 CC EMBL: X89974; CAA61893.1; JOINED.
 CC EMBL: X89975; CAA61893.1; JOINED.
 CC EMBL: X89976; CAA61893.1; JOINED.
 CC EMBL: X89977; CAA61893.1; JOINED.
 CC EMBL: X89978; CAA61893.1; JOINED.
 CC EMBL: X89979; CAA61893.1; JOINED.
 CC EMBL: X89980; CAA61893.1; JOINED.
 CC EMBL: X89981; CAA61893.1; JOINED.
 CC EMBL: X89982; CAA61893.1; JOINED.
 CC EMBL: X89983; CAA61893.1; JOINED.
 CC EMBL: X89984; CAA61893.1; JOINED.
 CC EMBL: X89985; CAA61893.1; JOINED.
 CC EMBL: X89986; CAA61893.1; JOINED.
 CC EMBL: X89987; CAA61893.1; JOINED.
 CC EMBL: X89988; CAA61893.1; JOINED.
 CC EMBL: X89989; CAA61893.1; JOINED.
 CC EMBL: X89990; CAA61893.1; JOINED.
 CC EMBL: X89991; CAA61893.1; JOINED.
 CC EMBL: X89992; CAA61893.1; JOINED.
 CC EMBL: X89993; CAA61893.1; JOINED.
 CC EMBL: X89994; CAA61893.1; JOINED.
 CC EMBL: X89995; CAA61893.1; JOINED.
 CC EMBL: X89996; CAA61893.1; JOINED.
 CC EMBL: X89997; CAA61893.1; JOINED.
 CC EMBL: X89998; CAA61893.1; JOINED.
 CC EMBL: X89999; CAA61893.1; JOINED.
 CC EMBL: X89900; CAA61893.1; JOINED.

RA	Komm B., Mohn K.;
RT	"Regulation of alpha 2(I), alpha 1(III), and alpha 2(V) collagen mRNAs by estradiol in the immature rat uterus";
RL	DNA 7:347-354 (1988).
CC	-1- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES ALONG WITH TYPE I COLLAGEN.
CC	-1- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE ALSO CROSS-LINKED VIA HYDROXYLINES.
CC	-1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL: X70369; CAA49832.1; .
DR	EMBL; AJ005395; CAA06510.1; .
DR	EMBL; M21354; AA040942.1; .
PIR	A23905; A23905.
PIR	SA1067; S41067.
DR	InterPro: IPR000087; Collagen.
DR	InterPro: IPR000885; Fib_collagen_C.
DR	InterPro: IPR001007; WFPC.
DR	PFAM; PF01410; COLFI_1.
DR	PFAM; PF01391; Collagen_6.
DR	PRODOM; PD002078; Fib_collagen_C; 1.
DR	SMART; SM00038; COLFI_1.
DR	PROSITE; PS01208; WFPC; PARTIAL.
KW	Extracellular matrix; Connective tissue; Repeat; Hydroxylation; Glycoprotein.
FT	NON_TER 1 1
FT	CHAIN 1 375
FT	PROPEP 376 636
FT	DOMAIN 368 368
FT	DISULFID 369 368
FT	DISULFID 369 369
FT	CONFLICT 340 429
FT	SEQUENCE 636 AA; 63332 MW; 61A48159FOLD01EE CRC64;
FT	1 1
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT	1 429
FT	1 63332
FT	1 61A48159FOLD01EE
FT	1 375
FT	1 636
FT	1 368
FT	1 369
FT	1 340
FT</td	

DR	EMBL; M18933; AAA37338_1; -.	RT	from type I procollagen and conservation of codon preferences.";
DR	EMBL; K03037; -; NOT_ANNOTATED_CDS.	RL	Biochem. J. 260:509-516(1989).
DR	EMBL; AK019448; BA31374_1; -.	RN	[2]
DR	EMBL; X57983; CAA41048_1; -.	RP	SEQUENCE OF 149-1225 FROM N.A.
DR	PIR; A22287; A22287.	RX	Medline=8938015; PubMed=2780304
DR	PIR; A27353; A27353.	RA	Janczko R.A., Ramirez F.;
DR	PIR; S16373; S16373.	RA	"Nucleotide and amino acid sequences of the entire human alpha 1 (III) collagen.";
DR	MGI; MGI_88453; Col3al.	RT	Nucleic Acids Res. 17:6742-6742(1989).
DR	InterPro; IPR000087; collagen.	RL	Nucleic Acids Res. 17:6742-6742(1989).
DR	InterPro; IPR000885; fib_collagen_C.	RN	[3]
DR	InterPro; IPR001007; WIFC.	RP	SEQUENCE OF 168-398.
DR	Pfam; PF01410; COLFI_1.	RX	Medline=77134724; PubMed=557335;
DR	Pfam; PF01391; Collagen_17.	RA	Sever J.M., Kang A.H.;
DR	PROSITE; PS01408; fib_collagen_C; 1.	RA	"Covalent structure of collagen: amino acid sequence of cyanogen bromide peptides from the amino-terminal segment of type III collagen of human liver.";
DR	SMART; SM00038; COLFI_1.	RT	RT
DR	SMART; SM00214; VNC_1.	RT	RT
KW	Extracellular matrix; Connective tissue; Repeat; Hydroxylation;	RL	RT
KW	Glycoprotein; Collagen; Signal;	RN	[4]
FT	SIGNAL 1 23 BY SIMILARITY.	RP	REVISIONS.
FT	PROPEP 24 154 AMINO-TERMINAL PROPEPTIDE.	RA	Sever J.M.;
FT	CHAIN 155 1203 COLLAGEN ALPHA 1(III) CHAIN.	RL	Submitted (DEC-1977) to the PIR data bank.
FT	PROPEP 1204 1464 CARBOXYL-TERMINAL PROPEPTIDE.	RN	[5]
FT	DOMAIN 31 90 VIFC.	RP	SEQUENCE OF 399-727.
FT	DOMAIN 155 169 NONHELICAL REGION (N-TERMINAL).	RX	Medline=79000343; PubMed=687591;
FT	DOMAIN 170 1195 TRIPLE-HELICAL REGION.	RA	Sever J.M., Kang A.H.;
FT	DOMAIN 1196 1464 NONHELICAL REGION (C-TERMINAL).	RA	"Covalent structure of collagen: amino acid sequence of five consecutive CNBr peptides from type III collagen of human liver.";
FT	CARBOHYD 262 262 O-LINKED (GAL-). (BY SIMILARITY).	RT	RT
FT	MOD_RES 262 262 HYDROXYLATION (BY SIMILARITY).	RL	RT
FT	MOD_RES 283 283 HYDROXYLATION (BY SIMILARITY).	RN	SEQUENCE OF 728-964.
FT	MOD_RES 859 859 HYDROXYLATION (BY SIMILARITY).	RP	Medline=80198282; PubMed=6246925;
FT	MOD_RES 976 976 HYDROXYLATION (BY SIMILARITY).	RX	Medline=80198282; PubMed=6246925;
FT	MOD_RES 1093 1093 HYDROXYLATION (BY SIMILARITY).	RA	Sever J.M., Mainardi C., Kang A.H.;
FT	MOD_RES 1105 1105 HYDROXYLATION (BY SIMILARITY).	RA	"Covalent structure of collagen: amino acid sequence of alpha 1 (III)-C5 from type III collagen of human liver.";
FT	DISULFID 1195 1195 INTERCHAIN (BY SIMILARITY).	RT	RT
FT	DISULFID 1196 1196 INTERCHAIN (BY SIMILARITY).	RL	Biochemistry 19:1583-1589(1980).
SQ	SEQUENCE 1464 AA; 138944 MW; 2104EC27A886090B CRC64;	RN	[7]
Query Match 9.3% Score 99; DB 1; Length 1464;	RP	SEQUENCE OF 950-1466 FROM N.A.	
Best Local Similarity 35.4% Pred. No. 6.5%; Mismatches 8; Indels 14; Gaps 4;	RA	Medline=88188827; PubMed=3337782;	
Matches 29; Conservative 8; Mismatches 31; Indels 14; Gaps 4;	RL	Markoo B.S., Dagleish R.;	
	RN	"Human pro alpha 1(III) collagen: cDNA sequence for the 3' end.";	
	RP	Nucleic Acids Res. 16:2337-2337(1988).	
QY 5 PDSVSECDREGRLQPAER--PPQLRGPAP-TSLOTEPQGNP-----EGNHGGEG 52	RX	[8]	
QY 972 PQGIKGESKPKGASGHNGRGPQGPQGPQGTAGEGRDNGPQSGPQGK 1031	RA	SEQUENCE OF 1065-1466 FROM N.A.	
DB 53 DSCPHGSPQGP LAP-PASP GP 72	RA	Medline=85087944; PubMed=6096827;	
QY 1032 DRGENGSPQGPAGPQGPQGP 1053	RL	Nucleic Acids Res. 16:11833-11833(1988).	
DB Homo sapiens (Human).	RA	Loedl H.R., Brinker J.M., May M., Pihlajaniemi T., Morrow S.,	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	RA	"Molecular cloning and carboxyl-propeptide analysis of human type III procollagen.";	
OC NCBI_TaxID: 9606;	RT	RT	
OX NCBI_TaxID: 9606;	RL	Rosenblom J., Myers J.C.;	
RN [1]	RN	"Molecular cloning and carboxyl-propeptide analysis of human type III procollagen.";	
RP SEQUENCE FROM N.A.	RA	RT	
RC TISSUE=Skin fibroblast;	RA	SEQUENCE OF 1176-1466 FROM N.A.	
RX MEDLINE=9350838; PubMed=2764886;	RA	Medline=8515600; PubMed=2579949;	
RX Ala_Kokko L., Kontusaari S., Baldwin C.T., Kuitaniemi H.,	RA	Chu M.-L., Wei D., de Wet W.J., Bernard M.P., Ramirez F.;	
RA Prockop D.J.;	RT	"Isolation of cDNA and genomic clones encoding human pro-alpha 1 (III) collagen. Partial characterization of the 3' end region of the gene.";	
RT "Structure of cDNA clones coding for the entire prepro alpha 1 (III) chain of human type III procollagen. Differences in protein structure	RT	RL	
RN [12]	RN	SEQUENCE OF 1161-1200 FROM N.A.	
RP MEDLINE=86187804; PubMed=3754462;	RA	Medline=86187804; PubMed=3754462;	

RA Miskulin M., Dagleish R., Klueve-Beckerman B., Rennard S.I., Tolstoshev P., Brantly M., Crystal R.G.; "Human type III collagen gene expression is coordinately modulated with the type I collagen genes during fibroblast growth."; *RL Biochemistry* 25:1408-1413(1986). [13]

RN SEQUENCE OF 1-170 FROM N.A.

RC TISSUE-Placenta; RX MEDLINE=88303360; PubMed=3405773;

RA Toman D., Ricca G., de Crombrughe B.; "Nucleotide sequence of a cDNA coding for the amino-terminal region of human propro alpha 1(III) collagen."; *Nucleic Acids Res.* 16:7201-7201(1988). [14]

RN SEQUENCE OF 1-176 FROM N.A.

RX MEDLINE=89378752; PubMed=2777083;

RA Benson-Chanda V., Su M.W., Weil D., Chu M.-L., Ramirez F.; "Cloning and analysis of the 5' portion of the human type-III procollagen gene (COL3A1)"; *Nucleic Acids Res.* 17:255-265(1989). [15]

RN REVIEW ON VARIANTs.

RA Kuivaniemi H., Tramp G., Prockop D.J.; "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-associated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels."; *RL Hum. Mutat.* 9:300-315(1997). [16]

RN VARIANT AORTIC ANEURYSM ARG-303, AND VARIANT THR-668.

RA Tramp G., Wu Y., Prockop D.J., Madhatheri S.L., Kleinert C., Earley J.J., Zhuang J., Noerregaard O., Darling R.C., Abbott W.M., Cole C.W., Jaakkola P., Reynaen M., Pearce W.H., Yao J.S.T., Rajamaa K., Smulders S.N., Gatalica Z., Ferrell R.E., Jimenez S.A., Jackson C.E., Michels V.V., Kaye M., Kuivaniemi H.; "Sequencing of cDNA from 50 unrelated patients reveals that mutations in the triple-helical domain of type III procollagen are an infrequent cause of aortic aneurysms."; *RL J. Clin. Invest.* 91:2539-2545(1993). [17]

RN VARIANT THR-698.

RA Zafarullah K., Kleinert C., Tramp G., Kuivaniemi H., Kontusaari S., Wu Y., Ganguly A., Prockop D.J.; "G to A polymorphism in exon 31 of the COL3A1 gene."; *RL Nucleic Acids Res.* 18:6180-6180(1990). [18]

RN VARIANT AORTIC ANEURYSM ARG-786.

RA Kontusaari S., Tramp G., Kuivaniemi H., Romanic A.M., Prockop D.J.; "A mutation in the gene for type III procollagen (COL3A1) in a family with aortic aneurysms."; *RL J. Clin. Invest.* 86:1465-1473(1990). [19]

RN VARIANT EDS-IV ARG-828.

RA MEDLINE=94016385; PubMed=8411057;

RA Richards A.J., Narcisi P., Lloyd J.C., Ferguson C., Pope F.M.; "The substitution of glycine 661 by arginine in type III collagen produces mutant molecules with different thermal stabilities and causes Ehlers-Danlos syndrome type IV."; *RL J. Med. Genet.* 30:690-693(1993). [20]

RN VARIANT EDS-IV ARG-957.

RA Tramp G., Kuivaniemi H., Shikata H., Prockop D.J.; "A single base mutation that substitutes serine for glycine 790 of the alpha 1 (III) chain of type III procollagen exposes an arginine and causes Ehlers-Danlos syndrome IV."; *RL J. Biol. Chem.* 264:1349-1352(1989). [21]

RN VARIANT EDS-IV VAL-960.

RA MEDLINE=95268429; PubMed=7749417;

RA Tramp G., de Paepe A., Nuytinck L., Madhatheri S.L., Kuivaniemi H.; "Substitution of valine for glycine 793 in type III procollagen in Ehlers-Danlos syndrome type IV."; *RL Hum. Mutat.* 5:179-181(1995). [22]

RN VARIANT EDS-IV GLU-1014.

RA MEDLINE=92316511; PubMed=1352273;

RA Richards A.J., Ward P.N., Narcisi P., Nicholls A.C., Lloyd J.C., Pope F.M.; "A single base mutation in the gene for type III collagen (COL3A1) converts glycine 847 to glutamic acid in a family with Ehlers-Danlos syndrome type IV. An unaffected family member is mosaic for the mutation."; *RL Hum. Genet.* 89:414-418(1992). [23]

RN VARIANT EDS-IV ASP-1050.

RA MEDLINE=90037070; PubMed=2808425;

RA Tramp G., Kuivaniemi H., Stolle C.A., Pope F.M., Prockop D.J.; "Single base mutation in the type III procollagen gene that converts the codon for glycine 83 to aspartate in a mild variant of Ehlers-Danlos syndrome IV."; *RL J. Biol. Chem.* 264:1913-1917(1989). [24]

RN VARIANT EDS-IV VAL-1077.

RA MEDLINE=91374480; PubMed=1895316;

RA Richards A.J., Lloyd J.C., Ward P.N., de Paepe A., Narcisi P., Pope F.M.; "Characterisation of a glycine to valine substitution at amino acid position 910 of the triple helical region of type III collagen in a patient with Ehlers-Danlos syndrome type IV."; *RL J. Med. Genet.* 28:458-463(1991). [25]

RN VARIANT EDS-IV GLU-1173.

RA MEDLINE=9302543; PubMed=1357232;

RA Johnson P.H., Richards A.J., Pope F.M., Hopkinson D.A.; "Query Match Score 99; DB 1; Length 1466; Best Local Similarity 27.8%; Pred. No. 6.5; Matches 37; Conservative 7; Mismatches 45; Indels 44; Gaps 5;

QY 5 PSDVSSCDREGQLOPAERPPQLPQAPSQQT--EP--OGNP----EGNHGEG 52

DB 973 PQGVKGSGKPGANGLSGERGPQGPQGPGLPLGLAGTGGPGRDGNPGSDGLPGRDGSQGKG 1032

QY 53 DSCPQHGSQGPQPLAP--PASPQPFATRSPFLFMRRSSLSRSSGQYSFDTDRSPAMSC 110

DB 1033 DRGENGSPGAPAPGHGPQGPV-----

QY 111 DKSTQTPSPPPCOA 123

DB 1063 DRGESEGPAGPAGA 1075

RESULT 11

COIL_HSVS7

ID COIL_HSVS7 STANDARD; PRT; 105 AA.

AC P5050;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Collagen-like protein

OS Herpesvirus saimiri (strain 484-77).

OC Gammaherpesviridae; Rhadinovirus.

OX NCBI_TaxID=10382;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90279084; PubMed=2161952;

RA Geck P., Whitaker S.A., Medveczky M.M., Medveczky P.G.; "Expression of collagenlike sequences by a tumor virus, herpesvirus saimiri"; *J. Virol.* 64:3509-3515(1990). [2]

ARE PRODUCED BY ALTERNATIVE SPLICING.																	
-1 - DOMAIN: ASN AT POSITION 47 OF THE HOMEobox MAY PARTICIPATE IN																	
CC -1 - REGULATING DNA BINDING ACTIVITY BY PROMOTING HOMO- AND																	
CC -1 - HETERO-DIMERIZATION.																	
CC -1 - SIMILARITY: CONTAINS 3 CUT DOMAINS.																	
CC -1 - SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEBOX PROTEINS.																	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).																	
CC -----																	
CC DR EMBL: M47499; -; NOT_ANNOTATED_CDS.																	
CC DR EMBL: AF047825; AAC78778.1; -.																	
CC DR HSSP; PI0037; IAUT.																	
CC DR TRANSFAC; T00100; -.																	
CC DR MIM: 116896; -.																	
CC DR InterPro; IPR033550; CUT [*] .																	
CC DR InterPro; IPR01356; Homeobox.																	
CC DR Pfam: PF002376; CUT; 3.																	
CC DR SMART; SM00385; HOX; 1.																	
CC DR PROSITE; PS00027; HOMEBOX_1; 1.																	
CC DR PROSITE; PS50071; HOMEBOX_2; 1.																	
CC KW Transcription regulation; Homeobox; DNA-binding; Repressor;																	
CC KW Developmental protein; Nuclear protein; Repeat; Repressor;																	
CC KW Colined coil; Alternative splicing; COILED_COIL (POTENTIAL).																	
FT DOMAIN 7 363 COILED_COIL (POTENTIAL).																	
FT DNA_BIND 542 629 CUT 1.																	
FT DOMAIN 669 725 COILED_COIL (POTENTIAL).																	
FT DNA_BIND 934 1021 CUT 2.																	
FT DNA_BIND 1117 1204 CUT 3.																	
FT DNA_BIND 1244 1303 HOMEBOX.																	
FT VARBIND 632 653 MISSING (IN ISOFORM 2).																	
SQ SEQUENCE 1505 AA; 164353 MW; 860E14D509D4DE11 CRC64;																	
Query Match 9.28; Score 97.5; DB 1; Length 1505;																	
Best Local Similarity 23.28; Pred. No. 8.6;																	
Matches 52; Conservative 25; Mismatches 82; Indels 65; Gaps 10;																	
QY 6 SDYSSCDEPREGRLQPAEPQLRPGAPPSLQTEPQGNPENHGGEDSPCHGSPQ --- 61																	
Db 404 SDLSGSARRKGKD-QPESRRGSPSLAPPSPQLPRNPQEOASNTNGTHQFSPAGLSQDFFS 462																	
QY 62 ---GPLAPAPASCPGPFAATPSPLFLFIMRR---SSLLSRS-----SGYFSFD---TD 102																	
Db 463 SSSLASPLASLKGKFLANS---LLQRQLMQSFSYSKAMQEAGSTSMSIFSTGPYSTNSTS 519																	
QY 103 RSPAPMSCLDKSTQSTOPSSP-----COAAGS----- 129																	
Db 520 QSPLQQSPVQNGMAPSPSCSESAGVSECEEMDTAEIARQVKQELRIKHNIGQRIFGHIVL 579																	
QY 130 AMASMROQAFPADMRPEIWIQAELRIGDEFNAYARRYFLNNYQ 173																	
Db 580 GLSQGSVSVL-ARPKPANKLTVR-GKE-PPHMKMQLFSDEQ 618																	
RESULT 1.3																	
DLG2_RAT 1D DLG2_RAT STANDARD; PRT; 852 AA.																	
AC Q63222; Q62939; P0548;																	
DT 01-NOV-1997 (Rel. 35, Created)																	
DT 01-NOV-1997 (Rel. 35, Last sequence update)																	
DT 16-OCT-2001 (Rel. 40, Last annotation update)																	
DE Channel associated protein of synapse-110 (Chapsyn-110) (Synaptic density protein PSD-95) (Discs, large homolog 2).																	
GN DLG2.																	
OS Rat; norvegicus (Rat);																	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																	

OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	FT	CONFLICT	626	641	MISSING (IN REF. 2).
OX	NCBI_TAXID=10116;	FT	CONFLICT	639	639	K -> A (IN REF. 3).
RN	[1]	FT	CONFLICT	726	726	F -> L (IN REF. 1).
RP	SEQUENCE FROM N.A.	FT	CONFLICT	733	733	N -> Y (IN REF. 2).
RA	Medline=9631088; PubMed=8755482;	FT	CONFLICT	749	749	E -> V (IN REF. 1).
RA	Kim E., Cho K.O., Rothschild A., Sheng M.;	FT	CONFLICT	756	756	L -> H (IN REF. 2).
RA	"Heteromultimerization and NMDA receptor-clustering activity of	FT	CONFLICT	791	792	KR -> NG (IN REF. 2).
RA	protein PSD-95 and alpha-syntrophin mediated by PDZ domains.";	FT	CONFLICT	794	794	T -> M (IN REF. 2).
RL	Chapron 110, a member of the PSD-95 family of proteins.;	SQ	SEQUENCE	852 AA;	94934 MW;	RD414AB9CF5B09 CRC64;
RN	[2]					
RP	SEQUENCE FROM N.A.					
RA	Medline=96193770; PubMed=8625413;					
RA	Brennan J.E., Chao D.S., Gee S.H., McGee A.W., Craven S.E.,					
RA	Santillano D.R., Wu Z., Huang F., Xia H., Peters M.F.,					
RA	Proehl S.C., Bredt D.S.;					
RA	"Interaction of nitric oxide synthase with the postsynaptic density					
RA	protein PSD-95 and alpha-syntrophin mediated by PDZ domains.";					
RL	Cell 84:757-767 (1996).					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RA	Irie M., Hata Y., Takai Y.;					
RL	Submitted (SEP-1996) to the EMBL/GenBank/DBJU databases					
CC	"- FUNCTION: INTERACTS WITH THE CYTOSPLASMIC TAIL OF NMDA RECEPTOR					
CC	SUBUNITS AS WELL AS POTASSIUM CHANNELS.					
CC	"- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.					
CC	"- SIMILARITY: CONTAINS 1 SH3 DOMAIN.					
CC	"- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN					
CC	"- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license (agreement) (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).					
CC						
DR	U49049; AB53243.1;					RESULT 14
EMBL	U50717; AAC52643.1;					DIG2_HUMAN
DR	HSSP; Q12959; 1PDR.					ID DIG2_HUMAN
DR	InterPro; IPR00619; Guanylate_kin.					STANDARD;
DR	InterPro; IPR001478; PDZ.					PRT;
DR	InterPro; IPR001452; SH3.					870 AA.
DR	PFam; PF00625; Guanylate_kin. 1.					
DR	PFam; PF00595; PDZ; 3.					
DR	SMART; SM00018; SH3; 1.					
DR	SMART; SM00072; GUKC; 1.					
DR	SMART; SM00228; PDZ; 3.					
DR	PROSITE; PS50052; GUANYLATE_KINASE_1; 1.					
DR	PROSITE; PS50106; PDZ; 3.					
DR	PROSITE; PS50002; SH3; 1.					
KW	SH3 domain; Repeat.					
FT	DOMAIN 98 184					CC SUBUNITS AS WELL AS POTASSIUM CHANNELS.
FT	DOMAIN 193 279					CC -!- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
FT	DOMAIN 421 501					CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
FT	DOMAIN 536 606					CC -!- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
FT	DOMAIN 662 852					
FT	CONFFLICT 181 182					This SWISS-PROT entry is copyright. It is produced through a collaboration
FT	CONFFLICT 228 228					between the Swiss Institute of Bioinformatics and the EMBL outstation -
FT	CONFFLICT 326 326					the European Bioinformatics Institute. There are no restrictions on its
FT	CONFFLICT 339 339					use by non profit institutions as long as its content is in no way
FT	CONFFLICT 450 454					modified and this statement is not removed. Usage by and for commercial
FT	CONFFLICT 464 465					entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
FT	CONFFLICT 474 474					
FT	CONFFLICT 476 476					EMBL; U32376; AA804949.1; -.
FT	CONFFLICT 478 478					DR HSSP; Q12959; 1PDR.
FT	CONFFLICT 484 486					DR MIM; 603583; -.
FT	CONFFLICT 506 506					DR InterPro; IPR000619; Guanylate_kin.
FT	CONFFLICT 569 569					DR InterPro; IPR001478; PDZ.
FT	CONFFLICT 586 586					

DR EMBL; L22545; AAA19787.1; -
 DR PDB; 1KOE; 16-FEB 99.
 DR MGD; MG1:88451; Coll8a1.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR011791; Laminin_G.
 DR InterPro; IPR0033129; TSPN.
 DR Pfam; PF01391; Collagen; 8.
 DR Pfam; PF02210; TSPN; 2.
 DR SMART; SM00382; LamG; 1.
 DR SMART; SM00210; TSPN; 1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
 KW 3D-structure.

FT SIGNAL 1 26
 FT CHAIN 27 26
 FT CHAIN 1344 1527
 FT DOMAIN 27 538
 FT DOMAIN 539 565
 FT DOMAIN 566 575
 FT DOMAIN 576 649
 FT DOMAIN 650 673
 FT DOMAIN 674 795
 FT DOMAIN 796 818
 FT DOMAIN 819 901
 FT DOMAIN 902 915
 FT DOMAIN 916 957
 FT DOMAIN 958 970
 FT DOMAIN 971 1043
 FT DOMAIN 1044 1053
 FT DOMAIN 1054 1086
 FT DOMAIN 1087 1098
 FT DOMAIN 1099 1122
 FT DOMAIN 1123 1129
 FT DOMAIN 1130 1181
 FT DOMAIN 1182 1194
 FT DOMAIN 1195 1212
 FT DOMAIN 1213 1527
 FT CARBOHYD 338 338
 FT CARBOHYD 700 700
 FT DISULFID 1376 1516
 FT DISULFID 1478 1508
 FT SITE 1104 1106
 FT VARSPLIC 1 212
 FT VARSPLIC 213 238
 FT CONFLICT 900 907
 FT CONFLICT 947 947
 FT CONFLICT 964 964
 FT CONFLICT 1157 1157
 FT CONFLICT 1266 1266
 FT CONFLICT 1276 1276
 FT CONFLICT 1437 1437
 FT SEQUENCE 1527 AA; 156008 MW; 9645045AF140B513 CRC64;

POTENTIAL.
 COLLAGEN ALPHA 1(XVIII) CHAIN.
 ENDOSTAVIN.
 NONHELCIAL REGION 1 (NC1).
 TRIPLE-HELICAL REGION 1 (COL1).
 NONHELCIAL REGION 2 (NC2).
 TRIPLE-HELICAL REGION 2 (COL2).
 NONHELCIAL REGION 3 (NC3).
 TRIPLE-HELICAL REGION 3 (COL3).
 NONHELCIAL REGION 4 (NC4).
 TRIPLE-HELICAL REGION 4 (COL4).
 NONHELCIAL REGION 5 (NC5).
 TRIPLE-HELICAL REGION 5 (COL5).
 NONHELCIAL REGION 6 (NC6).
 TRIPLE-HELICAL REGION 6 (COL6).
 NONHELCIAL REGION 7 (NC7).
 TRIPLE-HELICAL REGION 7 (COL7).
 NONHELCIAL REGION 8 (NC8).
 TRIPLE-HELICAL REGION 8 (COL8).
 NONHELCIAL REGION 9 (NC9).
 TRIPLE-HELICAL REGION 9 (COL9).
 NONHELCIAL REGION 10 (NC10).
 TRIPLE-HELICAL REGION 10 (COL10).
 NONHELCIAL REGION 11 (NC11).
 N-LINKED (GLCNAC . . .) (POENTIAL).
 N-LINKED (GLCNAC . . .) (POENTIAL).

CELL ATTACHMENT SITE (POENTIAL).
 MISSING (IN SHORT ISOFORM).
 AYPTQLPPFQSNLQAPGRPSAPPDF -> MAPRWHLIDVL
 TSIVLVLVARYVSAE (IN SHORT ISOFORM).

Query Match 9.08; Score 96; DB 1; Length 1527;
 Best Local Similarity 31.98; Pred. No. 11;
 Matches 36; Conservative 5; Mismatches 40; Indels 32; Gaps 6;

Qy 16 GROQPAERPPQLRGAPTSLOTEPQGNPNEGHHGGEDSCPHG-----SPQGPLAPPA 68
 Db 594 GRDGEPSDPGEDGRPG----DTGPQSFPP----GTPSDVPGKGEKDPGIGPRGPQPGP 644

Qy 69 SPQPFATRSPL-FIFMFRSLLRSRSSGGYFSDTDREPAAPMSCDK57QTPPP 120
 Db 645 PPGPSFRQDKLTFIDMEGSG-----FSGDIESLRGPPRG-----FPGP 682

THIS PAGE BLANK (USPTO)

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 10:01:02 ; Search time 26.66 Seconds (without alignments)

1284.809 Million cell updates/sec

Title: US-09-508-832-10

Perfect score: 1065

Sequence: 1 MAKOPSDVSSECDREGQLQ. PRMVILRLRYIVRLVWRMH 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_19:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp Rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_virus:
16: sp_bacteriophage:
17: sp_archaea:

ALIGNMENTS

RESULT 1
036424 PRELIMINARY;
ID 036424;
AC 036424;
DT 01-JAN-1998 (TREMBLE1. 05, Created)
DT 01-JAN-1998 (TREMBLE1. 05, Last sequence update)
DT 01-NOV-1998 (TREMBLE1. 08, Last annotation update)
DE PUTATIVE GLYCOPROTEIN.
OS Alcelaphine herpesvirus 1 (wildebeest herpesvirus).
OC Viruses; dsDNA viruses; no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhabdovirus.
OX NCBI_TAXID=35252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN="500";
RX MEDLINE=97404659; PubMed=9261371;
RA Enser A., Pilanz R., Fleckenstein B.;
RT J. Virol. 71:6517-6525 (1997).
RL AAC005370;
DR AF005370;
SQ SEQUENCE A72 AA; 50142 MW; 0DFFEF854764B8342 CRC54;

Query Match 10.4%; Score 110.5; DB 12; Length 472;
Best Local Similarity 22.9%; Pred. No. 0.0067;
Matches 48; Conservative 19; Mismatches 78; Indels 65; Gaps 7;

QY 25 PPQLRPQAPTSL-----QTEPOGNPEGNH-----48
Db 279 PPCTKPPCPTPLMDPCAVLSSSSTPPDRRKJTHRKVKVTGTPPVYLTISDDLD 338
O96424 alcelaphine herpes simplex
Q9nx00 homo sapien
Q9ve45 drosophila
Q96027 schizosaccharomyces pombe
Q9z58 streptomyces
Q9e200 streptomyces
Q95135 homo sapien
Q91zd2 mus musculus
Q96s97 homo sapien
Q91907 pagrus major
Q99893 human herpes simplex
Q9nf92 ascaris suum
Q9fcz5 mus musculus
Q96vz2 pneumocystis
Q9udu7 homo sapien

QY 49 --GGEGDSCPHGSPQPLAPPASPQPFATRSPLIFIMMRSSLLSRSRSGYFSDTDRSP 105
Db 339 DRKGAAAG-----GGPKPKPLPSRDPGSSGATPRSK 386
106 APNSCDKSTQTPPPCCQAFNHYSAMASMRQAPPADMRPEIWIAQELRRIGDEFNAYYA 164
Db 387 APAGRGRQEPPPLPP-----RDYDPASPEALRLRDIVQKVRYIGGAYNELEA 435

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	100.5	10.4	472	12	O96424		O96424 alcelaphine herpes simplex
2	106.5	10.0	620	12	P89435		Q9nx00 homo sapien
3	106.5	10.0	783	4	Q9NxD0		Q9ve45 drosophila
4	106.5	10.0	950	5	Q9Vp45		Q96027 schizosaccharomyces pombe
5	105.5	9.9	574	3	Q36027		Q9z58 streptomyces
6	103.5	9.7	556	2	Q9Zs8		Q9e200 streptomyces
7	103.5	9.7	556	2	Q9S20		Q95135 homo sapien
8	103.5	9.7	1051	4	Q95135		Q91zd2 mus musculus
9	102.5	9.6	1517	11	Q91zd2		Q96s97 homo sapien
10	102	9.6	321	4	Q96s07		Q91907 pagrus major
11	101	9.5	1820	13	Q97907		Q99893 human herpes simplex
12	100.5	9.4	620	12	Q89893		Q9nf92 ascaris suum
13	99.5	9.3	280	5	Q9nf92		Q9fcz5 mus musculus
14	99.5	9.3	378	11	Q9cSY5		Q96vz2 pneumocystis
15	99.5	9.3	874	3	Q96vz2		Q9udu7 homo sapien
16	99	9.3	257	4	Q9udu7		

Qy	165 RYVFLNNYQAAEDHPRMVILRLRLRYIVRLY 194	48 HGGEGDSCPQHGSPQGPAPPASPGPFATRSPLFLFMRRSSLLSRSSGGYFSFDTDRSPAP 107
Db	436 -1FDRNVEKQRERERMDYALIAJILLI 463	69 APGE---KARPASPALSEASSGPPPTDPL-----SPGGAHAIDOPDCSPP 111
RESULT 2		
P89435	PRELIMINARY ; PRT ; 620 AA.	
ID P89435 ; AC P89435 ;		
DT 01-MAY-1997 (TREMBLrel. 03, Created)	Qy 108 MSCD---KSTQTOPS--PP--CQAFNHYSAM 131	
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)	Db 112 PDPDPMMWSAIPNALPHILAEFFERHLRGL 143	
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE DEOXYRIBONUCLEASE.		
GN UL12.		
OS Herpes simplex virus (type 2).	RESULT 3	
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;	Q9NxD0 ID Q9NxD0; PRELIMINARY ; PRT ; 783 AA.	
OC Alphaherpesvirinae; Simplexvirus.	AC Q9NxD0;	
OX NCBI_TaxID=10310; [1]	DT 01-OCT-2000 (TREMBLrel. 15, Created)	
RN [1]	DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
RP SEQUENCE FROM N.A.	DE CDNA FLJ20315 FIS, CLONE HEP07873.	
RC STRAIN=HG52;	OS Homo sapiens (Human).	
RX MEDLINE=3027242; [2]	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Futeleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID=9606;	
RA McGeech D.J., Moss H.W., McNab D., Frame M.C.;	OX [1]	
RT "DNA sequence and genetic content of the HindIII 1 region in the short unique component of the herpes simplex virus type 2 genome: identification of the gene encoding glycoprotein G, and evolutionary comparisons."	RP SEQUENCE FROM N.A.	
RT J. Gen. Virol. 68:19-38(1987).	RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.; RT "NEEDO human cDNA sequencing project.";	
RN [2]	RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.	
RP SEQUENCE FROM N.A.	CC -! SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.	
RC STRAIN=HG52;	DR EMBL; AK000322; P28990; 1CNC.	
RX MEDLINE=30278430; [3]	DR HSSP; P28990; 1CNC.	
RA Everett R., Fenwick M.;	DR InterPro; IPR001841; Zinc_finger.	
RT "Comparative DNA sequence analysis of the host shut-off genes of different strains of herpes simplex virus: type 2 strain HG52 encodes a truncated UL41 product."	DR Pfam; PF00097; zf-C3HC4; 1.	
RT R. J. Gen. Virol. 71:1387-1390(1990).	DR SMART; SM00184; RING; 1.	
RN [3]	KW Zinc_finger. J. Gen. Virol. 71:1387-1390(1990).	
RP SEQUENCE FROM N.A.	SEQUENCE 783 AA; 85594 MW; 5845D826F0734FA5 CRC64;	
RC STRAIN=HG52;	Query Match 10 0%; Score 106.5; DB 4; Length 783;	
RX MEDLINE=32113549; [4]	Best Local Similarity 32.6%; Pred. No. 0.03; Mismatches 42; Conservative 11; Indels 17; Gaps 7;	
RA McGeech D.J., Cunningham C., McIntyre G., Dolan A.;	Qy 16 GRLQPAERPPOLRGPAP-TSLOQTEPGQNPENGNHGGEDSCSPHGG---SPOGPLA-PPASP 70	
RT "Comparative sequence analysis of the long repeat regions and adjoining parts of the long unique regions in the genomes of herpes simplex viruses types 1 and 2."	Db 566 GRKIGPETGVQSPRPIPTQPEPP-SPDQQVTSNSAAPSRLSINQCPRLPEAP 624	
RT J. Gen. Virol. 72:3057-3075(1991).	Qy 71 GPFATRSPLEIFMRSSLLSRSSGGYFSFDTDR-----SPABMSCKDSTQTPSPCPQA 123	
RN [4]	Db 625 GPVDASS--ICPSTSFLNQKSSLARHPQRKGGPSEPTPGSRPQDA-TVHPACQI 680	
RP SEQUENCE FROM N.A.	Qy 124 FNHYISAMA 132	
RC STRAIN=HG52;	Db 681 FPHXTPSVA 689	
RX MEDLINE=22356101; [5]	RESULT 4	
RA Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeech D.J.;	Q9VE45 ID Q9VE45; PRELIMINARY ; PRT ; 950 AA.	
RT A novel herpes simplex virus gene (UL49A) encodes a putative membrane protein with counterparts in other herpesviruses.";	AC Q9VE45;	
RT R. J. Gen. Virol. 73:2167-2171(1992).	DT 01-MAY-2000 (TREMBLrel. 13, Created)	
RN [5]	DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)	
RP SEQUENCE FROM N.A.	DE CG7709 PROTEIN.	
RC STRAIN=HG52;	GN CG7709.	
RA Dolan A.;	OS Drosophila melanogaster (Fruit fly).	
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.	DR EMBL; CAB0099; CAB0099; 72.1; -	
DR InterPro; IPR001616; Herpes_alk_exo.	AC Q9VE45;	
DR Pfam; PF01771; Herpes_alk_exo; 1.	DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DR PRINTS; PR00924; ALKEXNUCLASE.	DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)	
SQ SEQUENCE 620 AA; 66140 MW; E8AD3CF9A664163F CRC64;	DE CG7709.	
Query Match 10 0%; Score 106.5; DB 12; Length 620;	OS Drosophila melanogaster (Fruit fly).	
Best Local Similarity 27.6%; Pred. No. 0.023; Mismatches 42; Conservative 17; Indels 39; Gaps 8;	OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; NCBI_TaxID=7227;	
Qy 2 AKQPSDVSSSECDRGROLQP-----AERPPQLR--PPGAPTS-LQTEPGNPGCN 47	RN [1]	
Db 9 AKRPPADPARDPDSPPKRPNPSLTLATVFGPRPAPPRTSPGAFGSHWPQSPRQPDQG 68	RP SEQUENCE FROM N.A.	
●	RC STRAIN=BERKELEY;	

RX MEDLINE=20196006; PubMed=10731132;	GN WSP1 OR SPAC4F10.15C.
RA Adams M.D., Celikiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	OS Schizosaccharomyces pombe (Fission yeast).
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	OC Schizosaccharomyces pombe (Fission yeast).
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
RA Sutton G.R., Woertman J.R., Yandell M.D., Zhang Q., Chen L.X.,	OC Schizosaccharomyces pombe (Fission yeast).
RA Brandon R.C., Rogers Y.-H.C., Blazier R.G., Champe M., Preiffner B.D.,	OC Schizosaccharomyces pombe (Fission yeast).
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	OC Schizosaccharomyces pombe (Fission yeast).
RA Abril J.F., Agbeyani A., An H.-J., Andrews-Pflanckho C., Baldwin D.,	NCBI_TAXID=4896;
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	RN [1]
RA Beeson K.Y., Betos P.V., Herman B.P., Bhandari D., Brottier P.,	RP SEQUENCE FROM N.A.
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,	RC STRAIN=JS21;
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	RA Zankel T.C., OW D.W.;
RA Cherry J.M., Cawley S., Dahlie C., Davenport L.B., Davies P.,	RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RA de Pablo B., Delcher A., Deng Z., May A.D., Dew I., Dietz S.M.,	RN [2]
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	RP SEQUENCE FROM N.A.
RA Durbin K.J., Evangelista C.C., Ferreria S., Fleischmann W.,	RC STRAIN=972;
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,	RC Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,	CC -!- SIMILARITY: TO YEAST LAS17.
RA Hostin D., Houston K.A., Howland T.J., Herni M.-H., Ibegwam C.,	CC AF038575; AAB9287; 1; -.
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kentison J.A., Ketchum K.A.,	DR EMBL; Z98980; CAB11718; 1; -.
RA Kimmel B.E., Koira C.D., Kraft K., Kravitz S., Kulp D., Lai Z.,	DR InterPro; IPR000697; PanBP1_WASP.
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	DR InterPro; IPR001960; WH1.
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	DR InterPro; IPR003124; WH2.
RA Mukulov G., Mishina N.V., Moharry C., Morris J., Moshefi A.,	DR Pfam; PF00568; WH1; 1.
RA Mount S.M., Moy M., Murphy B., Morris J., Nelson D.L.,	DR Pfam; PF02205; WH2; 1.
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Paclob J.M.,	DR SMART; SM00461; WH1; 1.
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	FT DOMAIN 311 317 POLY-PRO.
RA Reinert K., Remmington K., Saunders R.D.C., Scheeler F., Shem H.,	FT DOMAIN 361 366 POLY-PRO.
RA Shue B.C., Siden-Kianos I., Simpson M., Skupsik M.P., Smith T.,	FT VARIANT 248 248 L->V (IN STRAIN JS21).
RA Spier E., Spradling A.C., Stapleton M., Strong M., Sun E.,	SQ SEQUENCE 574 AA: 59605 MW: C6E5EFC6A6A02F0E4 CRC44;
RA Svirsas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,	Query Match 9.98; Score 105.5; DB 3; Length 574;
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,	Best Local Similarity 26.8%; Pred. No. 0.027;
RA Williams S.M., Woodward T., Worley K.C., Wu D., Yang S., Yao Q.A.,	Matches 41; Conservative 17; Mismatches 66; Indels 29; Gaps 6;
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	QY 21 PAERPPOLRGPAPTSQUTEPQGNPEG--- NHGGEGDSCPHGSPQGLPAPPSPGP---- 72
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., zhu X., Smith H.O.,	Db 424 PPSAPSPPLPSAPSL--- PMGAPAAPPLPPSAPIAPPLPAPMAAPLPPAPAPAP 480
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;	QY 73 -PATRSPSLFIM-----R RSSLSSSSGYSFSFTD RSPAPMSCDKSTQTPSPPCQA- 123
RT "The genome sequence of <i>Drosophila melanogaster</i> .";	Db 481 APAPAAVASTAELPQDGRNLMASTRASGMDLRSRKVSASPVASTKTSNPPVEAP 540
RL Science 287:2185-2195(2000).	QY 124 -ENHYLSAMAS-----MRQAEPMRPEIW 147
DR EMBL; AE03723; AAF55384; 1; -.	Db 541 PSNNLMDALASLNQRTKVAQSDEEDEDDEW 573
DR FlyBase; FBgn0038642; CG7709	RN [1]
DR InterPro; IPR002965; P_rich_extensn.	RP SEQUENCE FROM N.A.
DR InterPro; IPR004019; YLP.	RC STRAIN=A3 (2);
DR Pfam; PF02757; YLP; 3	RA Bakal C.J., Davies J.E.;
DR PRINTS; PR01217; PRICHEXTNSN.	RT Cloning, nucleotide sequence and expression of a serine/threonine
SQ SEQUENCE 950 AA: 91280 MW: B57AB0EA4D30D3D7 CRC64;	RT protein kinase gene from <i>Streptomyces coelicolor</i> ";
Query Match 10.0%; Score 106.5; DB 5; Length 950;	AC DT 01-MAY-1999 (TREMBLrel. 10, Created)
Best Local Similarity 30.3%; Pred. No. 0.038; Mismatches 58; Indels 13; Gaps 4;	DB DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
Matches 37; Conservative 14; Mismatches 58; Indels 13; Gaps 4;	DE DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
QY 30 PGAPTSQUTEPQGNPEGNHGGEGDSCPHGSPQGLPAPPSPG----GPFATRSPLFIFMRSS 87	DE SERINE/THREONINE PROTEIN KINASE.
Db 785 PSAPSSYYAAPPSPSSSSSSGCGYPAFPNSNTSAPIAPSSSYGAPASGPSSFSAPSSYY 842	GN OS Streptomyces coelicolor.
QY 88 LLSRSSGGYFSFDTDSP-APMSCDKSTQTPSPPCOA----- NHYLSAMASMRQAE 138	OC Bacteria; Firmicutes; Actinobacteria; Actinomycetaceae; Streptomyces;
Db 843 GAPSTGSGSSSFSSSSSSGSSSSSSGSSSSSSGSSSSSSGSSSSSSGSSSSSSGSSSSSSGSS 902	OC Actinomycetales; Streptomyceae; Streptomyces.
QY 139 PA 140	OX NCBI_TAXID=1902;
Db 903 PA 904	RN
RESULT 5	RP
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	RC STRAIN=A3 (2);
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	RA Bakal C.J., Davies J.E.;
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	RT Cloning, nucleotide sequence and expression of a serine/threonine
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	RT protein kinase gene from <i>Streptomyces coelicolor</i> ";
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	CC -!- SIMILARITY: BEALC406.1; -.
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	DR EMBL; AF094711; AAC64406.1; -.
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	DR InterPro; IPR000719; Euk_pk kinase.
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	DR InterPro; IPR02290; Ser_thr_pk kinase.
036027 ID 036027; PRELIMINARY; PRT; 574 AA.	DE WISKOTT-ALDRICH SYNDROME PROTEIN HOMOLOG 1.

DR	Pfam; PF00069; Pkinase; 1. PROSTE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.	QY	1 MAKOPSD-----VSSEC DREGRQLOPAERPQLRGA PAPSLSQTEPQCNPEGN----HG 49	
DR	PROSTE; PS50011; PROTEIN_KINASE_DOM; 1.	Db	268 LKKNPNEE RTPSAEMMRD ECLRVAASFQAA----SCAGYQSAVFPV 320	
DR	PROSTE; PS00108; PROTEIN_KINASE_ST; 1.	QY	50 GEGDSCPHGSPQGPLAPAPSPGP-----ATRSPLFEMR RSSLLSRS SSSGYFSF DTD 102	
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.	Db	321 GOQTPAPG PVPQYQPTSPGPNPYGT PAPAAHSP AYGPQQ----AGY----QT 368	
SEQUENCE	556 AA; 58328 MW; B042F9E550870743 CRC64;	QY	103 RSPAPMSCDKSTQPSPPQFQAFNHYLSAMAS 133	
Query Match	9.78; Score 103.5; DB 2; Length 556;	Db	369 PAPAPYQ QQQAAATPPP----YNLTPSAQGS 395	
Best Local Similarity	27.8%; Pred. No. 0.041; Indels 41; Gaps 8;	RESULT	8	
Matches	42; Conservative 17; Mismatches 51; Indels 41; Gaps 8;	095135	PRELIMINARY; PRTR; 1051 AA.	
QY	1 MAKOPSD-----VSSEC DREGRQLOPAERPQLRGA PAPSLSQTEPQCNPEGN----HG 49	ID	095135 PRELIMINARY; PRTR; 1051 AA.	
Db	268 LKKNPNEE RTPSAEMMRD ECLRVAASFQAA----SCAGYQSAVFPV 320	AC	095135	
QY	50 GEGDSCPHGSPQGPLAPAPSPGP-----ATRSPLFEMR RSSLLSRS SSSGYFSF DTD 102	DT	01-MAY-1999 (TREMBLrel. 10, Created)	
Db	321 GOQTPAPG PVPQYQPTSPGPNPYGT PAPAAHSP AYGPQQ----AGY----QT 368	DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)	
QY	103 RSPAPMSCDKSTQPSPPQFQAFNHYLSAMAS 133	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
Db	369 PAPAPYQ QQQAAATPPP----YNLTPSAQGS 395	DE	ATAXIN-2-LIKE PROTEIN A2L P.	
RESULT	7	GN	A2LG.	
Q9S2C0	PRELIMINARY; PRTR; 556 AA.	OS	Homo sapiens (Human).	
AC	Q9S2C0; 01-MAY-2000 (TREMBLrel. 13, Created)	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	NCBI_TaxID	9606; Homo sapiens (Human).	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	RN	[1]	
DE	SERINE/THREONINE PROTEIN KINASE.	RP	SEQUENCE FROM N.A.	
GN	Streptomyces coelicolor.	RC	SEQUENCE=BRAIN;	
OS	Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;	RA	Xia J.-h., Liu C.-Y., Wang D.-a., Ruan Q.-g., Deng H.-x.;	
OC	Actinomycetales; Streptomyceae; Streptomyces; Streptomyces.	RT	"A splicing form of human ataxin-2 like gene obtained from adult brain".	
OX	NCBI_TaxID=1902;	RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.	
RN	SEQUENCE FROM N.A.	DR	EMBL; AF034373; AAC69607; 1.	
RC	STRAIN="A3(2);	DR	InterPro; IPR002965; P_rich_extrsn.	
RA	Saunders D.C., Harris D.; Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.	DR	PRINTS; PRO1217; PRICHEATENS.	
RN	[2]	SEQUENCE	1051 AA; 11048 MW; EBB21D51A17FD6B9 CRC64;	
RP	SEQUENCE FROM N.A.	Query	REGRQLQPAERPQLRGA PAPS-----LQTEPQGNPEGN HGGESDSCPHGSPQGPLAPAS 69	
RC	STRAIN="A3(2);	Match	9.7%; Score 103.5; DB 4; Length 1051;	
RA	James K.D., Parkhill J., Barrell B.G., Rajandream M.A.; Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.	Best_Local_Similarity	32.0%; Pred. No. 0.086; Mismatches 8; Indels 37; Gaps 6;	
RN	[3]	Matches	40; Conservative 8; Mismatches 40; Indels 37; Gaps 6;	
RP	SEQUENCE FROM N.A.	Db	7 RDAHFPPA RPPPLP RASSR GFLSPPP-----TPRS PRPTAGGC 52	
RC	STRAIN="A3(2);	QY	14 REGRQLQPAERPQLRGA PAPS-----LQTEPQGNPEGN HGGESDSCPHGSPQGPLAPAS 69	
RA	Redenbach M., Kieser H.M., Denapaitre D., Eichner A., Cullum J., Kinashi H., Hopwood D.A.; Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.	Db	70 PGPFATRS-----PLFIMR RSSLLSRS SSSG-----YFSF DTD RSPAPMSCDKSTQ 115	
RN	[4]	SEQUENCE FROM N.A.	Db	53 PIPLASRALQRGPSGP-----SPSLNSNSPGRCHIVEASAATTALPAPAAAPHATCR 107
RP	SEQUENCE FROM N.A.	QY	116 TPSPPP 120	
RC	STRAIN="A3(2);	Db	108 GPSAP 112	
RX	MEDLINE=9/00051; PubMed=8843436;	RESULT	9	
RA	Redenbach M., Kieser H.M., Denapaitre D., Eichner A., Cullum J., Kinashi H., Hopwood D.A.; Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.	ID	0912D2 PRELIMINARY; PRTR; 1517 AA.	
RN	[5]	SEQUENCE FROM N.A.	AC	Q912D2
RP	"A set of ordered cosmidns and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";	DT	01-DEC-2001 (TREMBLrel. 19, Created)	
RT	Mol. Microbiol. 21:77-96(1996).	DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
RT	C - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
RL	DR	DE	CCAAAT DISPLACEMENT PROTEIN CDP.	
CC	EMBL; AL096822; CAB46944; 1.	GN	CUTTJ.	
DR	InterPro; IPR000719; Euk_kinase.	OS	Mus musculus (Mouse).	
DR	InterPro; IPR00290; ser-thr_kinase.	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
DR	Pfam; PF00069; Pkinase; 1. PROTEIN_KINASE_ATP; UNKNOWN_1.	NCBI_TaxID	Muridae; Murinae; Mus.	
DR	PROSTE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.	RN	[1]	
DR	PROSTE; PS50011; PROTEIN_KINASE_DOM; 1.	RP	SEQUENCE FROM N.A.	
DR	PROSTE; PS00108; PROTEIN_KINASE_ST; 1.	RX	MEDLINE=21429086; PubMed=11544187;	
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.	EL113 T., Gambardella L., Horcher M., Tschanz S., Capol J.,		
SEQUENCE	556 AA; 58262 MW; D4A22A2E39F4ADC1 CRC64;	QY	Query Match 9.7%; Score 103.5; DB 2; Length 556;	
Best Local Similarity	27.8%; Pred. No. 0.041; Mismatches 51; Indels 41; Gaps 8;	Db	RP SEQUENCE FROM N.A.	
Matches	42; Conservative 17; Mismatches 51; Indels 41; Gaps 8;	RX	RX	

RA	Bertram P., Jochum W., Barrandon Y., Busslinger M.;	RESULT 11	Db 240 RTRVHTSAAFBALRKVPCYSYGQKLSKLAIRIACNYILSL 282
RT	"The transcriptional repressor CDP (Cut11) is essential for epithelial cell differentiation of the lung and the hair follicle.;"		
RL	Genes Dev 15:2307-2319 (2001).		
DR	EMBL: AY037807; AAK59386; 1;		
SQ	SEQUENCE 1517 AA; 165852 MW; B17A8740621EBBAC CRC64;		
Query Match 9.6%; Score 102.5; DB 11; Length 1517;			
Best Local Similarity 23.7%; Pred. No. 0.17; Indels 73; Gaps 9;			
Matches 54; Conservative 25; Mismatches 76;			
QY 6 SDSSECDREGLOQPAERPPQLRPGAPTSQQT--EPOGQPEGNHHGGEGDSCPHGSPQ 62	RESULT 12	Db 241 RTRVHTSAAFBALRKVPCYSYGQKLSKLAIRIACNYILSL 282	
Db 404 SDISGSARRKGQDQEPERRGPPLAPSPPPQPRNTPEQVSTNGTH-----HRSPAG 455	Q91907 PRELIMINARY; PRT; 1820 AA.	Q91907 PRELIMINARY; PRT; 1820 AA.	
QY 63 -----PLAPPASPGPFATRSPLFMRRS-----SLLSSSSGGFSSF 100	Q91907 (TREMBLrel. 15, Created)	Q91907 (TREMBLrel. 15, Last sequence update)	
Db 456 LSQDFESSNLASPLPLASTSKFALNSLQLRQMOSFYSKAMQEAGSTSTFSTGYSTN 515	DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 19, Last annotation update)	
QY 101 TDRSPADM--SCDKSTQTPSPPP-----CQAFN 125	GN COLV/XIA1.	DE PRO-ALPHA 1 TYPE V/XI COLLAGEN.	
Db 516 SISSPSPSLQQSPEDVNGMAPSPSQQSEASGSISSEGEELDTAEARQKEOLIKNIGORIFG 575	OS Pagrurus major (Red sea bream) (Chrysophrys major).	OC Actinopterygii; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthomorpha; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Perciformes; Percoidae; Pagrus.	
QY 126 HYLSAMASMRQAEPADMRPTEIWIAQELRIGIDDEFNAYAYARRFLNNYQ 173	OC Sparidae; Pagrus.	OC NCB1 TAXID=143350;	
Db 576 HYVLGLSQGSYSEITLARKPKNNKLTVR--GRE--PFHKMQLFSLDEQ 618	RP SEQUENCE FROM N.A.	RN	
RESULT 10	RP MEDLINE=21240220; PubMed=11342118;	RX	
Q96SQ7 PRELIMINARY; PRT; 321 AA.	RA Touhata K., Tanaka H., Yokoyama Y., Sakaguchi M., Toyohara H.;	RA Touhata K., Tanaka H., Yokoyama Y., Sakaguchi M., Toyohara H.;	
AC Q96SQ7 PRELIMINARY; PRT; 321 AA.	RT Structure of a full-length cDNA clone for the pro-1(V/XI) collagen chain of red seabream. "	RT Structure of a full-length cDNA clone for the pro-1(V/XI) collagen chain of red seabream. "	
DT 01-DEC-2001 (TREMBLrel. 19, Created)	RL Biochim. Biophys. Acta 1517:323-326(2001).	RL Biochim. Biophys. Acta 1517:323-326(2001).	
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)	DR InterPro: IPR000087; Collagen.	DR InterPro: IPR000087; Collagen.	
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)	DR InterPro: IPR001791; Laminin_G.	DR InterPro: IPR001791; Laminin_G.	
DE CDNA FEJ14708 FIS. CLONE M#2RP3000603, WEAKLY SIMILAR TO DE NEUROGENIC DIFFERENTIATION FACTOR 1.	DR InterPro: IPR001230; Frenin_Itn.	DR InterPro: IPR001230; Frenin_Itn.	
OS Homo sapiens (Human).	DR InterPro: IPR003129; TSPN.	DR InterPro: IPR003129; TSPN.	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	DR Pfam: PF02220; TSPN; 1.	DR Pfam: PF02220; TSPN; 1.	
OC NCBI_TAXID=9606;	DR Pfam: PF01410; COLIFI; 1.	DR Pfam: PF01410; COLIFI; 1.	
RN	RP SEQUENCE FROM N.A.	DR ProdDom: PDD02078; Fib_collagen_C; 1.	
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishizawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K., Nishizawa T., Kimata M., Watanabe M., Hiroka S., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahori K., Masuho Y., Kaneko K.;	DR SMART: SM00038; COLIFI; 1.	DR SMART: SM00038; COLIFI; 1.	
RT "NEDO human cDNA sequencing project.";	DR SMART: SM00082; LamG; 1.	DR SMART: SM00082; LamG; 1.	
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.	DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.	DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.	
DR EMBL: AK027614; BAB55233.1;	KW Collagen.	KW Collagen.	
SQ SEQUENCE 321 AA; 34627 MW; 8FB9C04C23BA4258 CRC64;	SQ SEQUENCE 1820 AA; 181678 MW; 46E45E8AF7AD3AE CRC64;	SQ SEQUENCE 1820 AA; 181678 MW; 46E45E8AF7AD3AE CRC64;	
Query Match 9.6%; Score 102; DB 4; Length 321;	Query Match 9.5%; Score 101; DB 13; Length 1820;		
Best Local Similarity 23.8%; Pred. No. 0.03; Indels 72; Gaps 9;	Best Local Similarity 29.4%; Pred. No. 0.29; Indels 54; Gaps 9;		
Matches 53; Conservative 21; Mismatches 77;	Matches 50; Conservative 8; Mismatches 58; Indels 54; Gaps 9;		
QY 2 AKQPSDVSECDREGLOQPAERPPQLRPGAPTSLOTEPQNPENHGGEDSCPHGSPQ 61	QY 30 PGAPTSLOTEPQNPENHGGEDSCPHGSPQPLAPPASPOPFPATSPLSPLFIMMRSSLL 89		
Db 101 APEVSDARKCFALGAVPGPLTPPPPPP----QSQAPGPEAQ----PFREP- 148	Db 1474 PGEPGERGDRGLPQGPQTTGGGGKGDGTCVGG-PQGPLGPQGPPG-----1SGPQGQKG 1523		
QY 62 GP-----LAPPASPGFATRSPLFMRRSLLSSSSGFSFDIQRSPMSCKSTQ 115	QY 90 SRSSSGYFSFSDTDRS----PAPMSCDSTQTPSPPP--CQAFNHYLSAMSRQAEPADM 142		
Db 149 GPPRPLLCAPPAPASA-----PPAPPAPPESTV 179	Db 1524 SKGSTGQKGDNGTIGPPGP-----PGPPGDIQPLPMQQSARKTRQAENQ-- 1572		
QY 116 TPSPPCQ-----AFNHYLSAMSRQ--AEPADMREWIQAERARRGDEF- 160	QY 143 RPEIWIQAELRRIGDEFNAYAA-----RNFV--LNNYQAAEDHPRM 181		
Db 180 RPAPPTRPGESSYSSHTVNNHQSSASPKRKGPEATAASSETRALQQTTRILLANARE 239	Db 1573 -----GDEMADYGVGDGVMEGMDDVFGSINNLK--QDERM 1607		
QY 161 -----AYARRYFLNNYQAOAEDHPRMTLRL-LRYVRL 193	Q989893 PRELIMINARY; PRT; 620 AA.	Q989893 PRELIMINARY; PRT; 620 AA.	
Db	DT 01-NOV-1996 (TREMBLrel. 01, Created)	AC Q989893 PRELIMINARY; PRT; 620 AA.	
	DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)	AC Q989893 PRELIMINARY; PRT; 620 AA.	
	DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)	AC Q989893 PRELIMINARY; PRT; 620 AA.	
	DE HYPOTHETICAL 67.1 KDA PROTEIN.	DE HYPOTHETICAL 67.1 KDA PROTEIN.	
	GN DR2.	GN DR2.	
	OS Human herpesvirus 6.	OS Human herpesvirus 6.	
	OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae.	OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae.	
	OC Betaherpesvirinae; Roseolovirus	OC Betaherpesvirinae; Roseolovirus	
	OX NCBI_TAXID=10368;	OX NCBI_TAXID=10368;	
	RN [1]	RN [1]	

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=5266321; PubMed=7747482;
RA Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J., Martin M.E., Efstathiou S., Craxton M., Macaulay H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution.";
RL Virology 209:29-51(1995).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=90080132; PubMed=2152817;
RA Lawrence G.L., Chee M., Craxton M.A., Gompels U.A., Honess R.W., Barrell B.G.;
RT "Human herpesvirus 6 is closely related to human cytomegalovirus.";
RL J. Virol. 64:287-299(1990).
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=91237802; PubMed=1851860;
RA Chang C.K., Balachandran N.;
RT "Identification, characterization, and sequence analysis of a cDNA encoding a phosphoprotein of human herpesvirus 6.";
RL J. Virol. 65:2884-2894(1991).
RN [4]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=1333007; PubMed=8305692;
RA Teo I.A., Griffin B.E., Jones M.D.;
RT "Characterization of the DNA polymerase gene of human herpesvirus 6.";
RL J. Virol. 65:4670-4680(1991).
RN [5]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=1651403;
RA Thomson B.J., Griffin B.E., Jones M.D.;
RT "Characterization of the DNA polymerase gene of human herpesvirus 6.";
RL J. Virol. 65:4670-4680(1991).
RN [6]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=91226542; PubMed=1851252;
RA Thomson B.J., Efstathiou S., Honess R.W.;
RT "Acquisition of the human adeno-associated virus type-2 rep gene by human herpesvirus type-6.";
RL Nature 351:78-80(1991).
RN [7]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=1374590; PubMed=1654446;
RA Martin M.E., Nicholas J., Thomson B.J., Newman C., Honess R.W.;
RT "Identification of a transactivating function mapping to the putative immediate-early locus of human herpesvirus 6.";
RL J. Virol. 65:5381-5390(1991).
RN [8]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=92333249; PubMed=1321206;
RA Efstathiou S.; Lawrence G.L., Brown C.M., Barrell B.C.;
RT "Identification of homologues to the human cytomegalovirus US22 gene family in human herpesvirus 6.";
RL J. Virol. 73:1661-1671(1992).
RN [9]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=12149942; PubMed=1310766;
RA Geng Y., Chaudhuri B., Josephs S.F., Wood C.;
RT "Identification and characterization of a human herpesvirus 6 gene segment that trans activates the human immunodeficiency virus type 1 promoter.";
RL J. Virol. 66:1564-1570(1992).
RN [10]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=9226671; PubMed=1374813;
RA Neipel F., Ellinger K., Fleckenstein B.;
RT "Gene for the major antigenic structural protein (p100) of human herpesvirus 6.";
RT "The right end of the unique region of the genome of human herpesvirus 6 U1102 contains a candidate immediate early gene enhancer and a homologue of the human cytomegalovirus US22 gene family.";
RL J. Gen. Virol. 66:3918-3924(1992).
RN [11]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=9233248; PubMed=1321205;
RA Thomson B.J., Honess R.W.;
RT "The right end of the unique region of the genome of human herpesvirus 6 U1102 contains a candidate immediate early gene enhancer and a homologue of the human cytomegalovirus US22 gene family.";
RL J. Gen. Virol. 66:3918-3924(1992).
RN [12]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=9318613; PubMed=8383182;
RA Ellinger K., Neipel F., Foa-Tomasz L., Campadelli-Fiume G.;
RT "The glycoprotein B homologue of human herpesvirus 6.";
RL J. Gen. Virol. 73:1649-1660(1992).
RN [13]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=93224882; PubMed=8305692;
RA Gompels U.A., Carrigan D.R., Carsi A.L., Arno J.;
RT "Two groups of human herpesvirus 6 identified by sequence analyses of laboratory strains and variants from Hodgkin's lymphoma and bone marrow transplant patients.";
RL J. Gen. Virol. 74:495-500(1993).
RN [14]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=9338439; PubMed=8397282;
RA Liu D.X., Gompels U.A., Nicholas J., Lelliott C.;
RT "Identification and expression of the human herpesvirus 6 glycoprotein H and interaction with an accessory 40K glycoprotein";
RL J. Gen. Virol. 74:1847-1857(1993).
RN [15]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=94025558; PubMed=7692666;
RA Liu D.X., Gompels U.A., Foa-Tomasz L., Campadelli-Fiume G.;
RT "Human herpesvirus 6 glycoprotein H and L homologs are components of the gp100 complex and the gH external domain is the target for neutralizing monoclonal antibodies.";
RL J. Virol. 197:12-22(1993).
RN [16]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=9333170; PubMed=7687803;
RA Pellett P., Sanchez-Martinez D., Dominguez G., Black J.B., Anton E., Greenamyer C., Dambach T.R.;
RT "A strongly immunoreactive virion protein of human herpesvirus 6 variant B strain Z29: identification and characterization of the gene and mapping of a variant-specific monoclonal antibody reactive epitope.";
RT "A strongly immunoreactive virion protein of human herpesvirus 6 variant B strain Z29: identification and characterization of the gene and mapping of a variant-specific monoclonal antibody reactive epitope.";
RL J. Virol. 195:521-531(1993).
RN [17]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=9332202; PubMed=7687301;
RA Pfeiffer B., Bernman Z.N., Neipel F., Chang C.K., Tirwatnapong S., Chandran B.;
RT "Identification and mapping of the gene encoding the glycoprotein gp82-gp105 of human herpesvirus 6 and mapping of the gp82-gp105 complex by monoclonal antibodies.";
RT "Infectivity determinants encoded in a conserved gene block of human herpesvirus-6.";
RT "Neutralizing epitope recognized by monoclonal antibodies.";
RL J. Virol. 67:4611-4620(1993).
RN [18]

RP SEQUENCE FROM N.A.

RC	STRAIN-U1102;	RT	Submitted;"
RX	PubMed=7844567;	RL	Submitted: AJ249273; CAB85466.1;
RA	Gompels U.A.; Macaulay H.A.;	DR	DR
PT	"Characterization of human telomeric repeat sequences from human	DR	InterPro; IPR00087; Collagen.
RT	herpesvirus 6 and relation to replication.;"	DR	InterPro; IPR002486; Col_cuticile_N.
RL	J. Gen. Virol. 76:451-458(1995).	DR	PFam; PF01391; Collagen; 2.
RN	[19]	DR	PFam; PF01484; Col_cuticile_N; 1.
RP	SEQUENCE FROM N.A.	KW	Collagen.
RC	STRAIN-U1102;	FT	NON_TER 1 1 1
RX	MEDLINE=94047392; PubMed=8230490;	SQ	SEQUENCE 280 AA; 27796 MW; E61086529408BC10 CRC64;
RA	Rewhurst S.; Dillard S.C.; Pellett P.E.; Dambaugh T.R.;	RT	Query Match 9.3%; Score 99.5%; DB 5; Length 280;
PT	"Identification of a lytic-c phase origin of DNA replication in human	Best Local Similarity 30.4%; Pred. No. 0.046;	
RT	herpesvirus 6B strain 229";	Matches 34; Conservative 9; Mismatches 36; Indels 33; Gaps 6;	
RL	J. Virol. 67:7680-7683(1993).	Qy	15 EGRLQDAERPPDOLPPAPTSLOTEPQNP - EGNHIGEGDSCPSPHGSPP - QGQPLAPPASP 70
RN	[20]	Db	193 EAGOLIGERPPPGPPGPPG - RPGPGEPPGPGPGEPPGPQGPQ 250
RP	SEQUENCE FROM N.A.	Qy	71 GPFATRSPFLFIMRSSLSSRSGYFSFTDRSPAPMSCDKSTQTPSPPCQ 122
RC	STRAIN-U1102;	Db	251 G-----QQGRPG----EPGQAGAPGSCDH----CPPAR 275
RX	Nicholas J.; Martin M.;	RESULT 14	
RA	"Nucleotide sequence analysis of a 38.5-kilobase-pair region of the	ID QCSY5 PRELIMINARY; PRT; 378 AA.	
PT	genome of human herpesvirus 6 encoding human cytomegalovirus	AC QCSY5;	
RT	immediate-early gene homologs and transactivating functions.";	DT 01-JUN-2001 (TREMBLrel. 17, Created)	
RL	J. Virol. 68:597-610(1994).	DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	
RN	[22]	DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
RP	SEQUENCE FROM N.A.	DE 261003413RICK PROTEIN (FRAGMENT).	
RC	STRAIN-U1102;	GN 261003413RICK.	
RX	MEDLINE=94202284; PubMed=8151768;	OS Mous musculus (Mouse).	
RA	Schiwe U.; Neipel F.; Schreiner D.; Fleckenstein B.;	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
PT	"Structure and transcription of an immediate-early region in the human	OC Mammalia; Etherilia; Rodentia; Sciurognathli; Muridae; Murinae; Mus.	
RT	herpesvirus 6 genome.";	NCBI_TaxID=10090;	
RL	J. Virol. 68:2978-2985(1994).	RN [1]	
RN	[23]	RP SEQUENCE FROM N.A.	
RP	SEQUENCE FROM N.A.	RC STRAIN-C57BL/6J; TISSUE="EMBRYO";	
RC	STRAIN-U1102;	RX MEDLINE=21056660; Published:11217851;	
RX	Thompson J.; Choudhury S.; Kashanchi F.; Berneman Z.,	RA Kawai J.; Shinagawa A.; Shibata K.; Yoshihro M.; Itoh M.; Ishii Y.;	
RA	Frenkel N.; Rosenthal L.J.;	RA Arakawa T.; Hira A.; Fukunishi Y.; Konno H.; Adachi J.; Fukuda S.;	
RA	Best Local Similarity 31.0%; Pred. No. 0.093; Gaps 5;	RA Aizawa K.; Izawa M.; Nishi K.; Miyosawa H.; Kondo S.; Yamanaoka I.;	
Matches 36; Conservative 10; Mismatches 51; Indels 19; Gaps 5;	RA Saito T.; Ohazaki Y.; Gojobori T.; Bono H.; Kasukawa T.; Saito R.;		
Qy	21 PAERPPOLRP-----GAPTSLOTEQGNPGNNG-----GEDDSCHPGSPOPLA 65	RA Radot A.; Matsuda H.A.; Ashburner M.; Batalov S.; Casavant T.;	
Db	486 PAEPPPTCRPRPRSSDTPLSAVSRSPAPVPPSTARVRFFLSSSSSSYSP-APL 544	RA Fleischmann W.; Gaasterland T.; Gissi C.; King B.; Kochiwa H.;	
Qy	66 PPSAPGPFATRSPFLFIMRSSLSSRSGYFSFTDRSPAPMSCDKSTQTPS 118	RA Kiehl P.; Lewis S.; Matsuo Y.; Nikaldo I.; Pesole G.; Quackenbush J.;	
Db	545 PPSPVSPSSPRSPFPIPRSPGLRAKPRVSSGHPAAFPSSAPARSERVTSPVS 600	RA Schriml L.M.; Staubli R.; Tomita R.; Wagner L.; Washio T.;	
RESULT 13		RA Sakai K.; Okido T.; Furuno M.; Aono H.; Baldarelli R.; Barsh G.;	
Q9NFZ9	PRELIMINARY;	RA Brown J.; Borrelli D.; Bojunga N.; Carninci P.; de Bonaldo M.F.;	
ID Q9NFZ9;	PRT; 280 AA.	RA Brownstein M.J.; Bult C.; Fletcher C.; Fujita M.; Garibaldi M.;	
AC Q9NFZ9;		RA Gustincich S.; Hill D.; Hofmann M.; Hume D.A.; Kamiya M.; Lee N.H.;	
DT 01-OCT-2000 (TREMBLrel. 15, Created)		RA Lyons P.; Marchionni L.; Mashima J.; Mazzarelli J.; Mombretti P.;	
DT 01-JUN-2001 (TREMBLrel. 15, Last sequence update)		RA Nordone P.; Ring B.; Ringwald M.; Rodriguez I.; Sakamoto N.;	
DE PUTATIVE CUTICULAR COLLAGEN (FRAGMENT).		RA Suzuki H.; Toyo-Oka K.; Wang K.H.; Weitz C.; Whittaker C.; Wilming L.;	
GN COL-D.		RA Wynshaw-Boris A.; Yoshida K.; Hasegawa Y.; Kawaji H.; Kohtsuki S.;	
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides) Ascaridida; Ascarididae;		RA Hayashizaki Y.;	
OC Ascarididae: Ascaris.		RA Functional annotation of a full-length mouse cDNA collection.";	
NCBI_TaxID=6253;		RT "Functional annotation of a full-length mouse cDNA collection.";	
RN [1]		RL Nature 409:685-690(2001).	
RP SEQUENCE FROM N.A.		DR MGI: MGI:19222863; BAB27764.1; -	
DI Mito C.; Betschart B.;		FT NON_TER 1 1 1	
RA "Identification of two collagen genes from the cuticle of Ascaris		SQ SEQUENCE 378 AA; 40207 MW; DDD783BA958F104C CRC64;	
Query Match 9.3%; Score 99.5%; DB 11; Length 378;		Qy 16 GRLQPAERPPQ -LRPGAPTSIQLQTEPQG-----NPEGNHGEGDSCPQG 62	

201	D-bb	GPILEQARPQNLLQLRAB--	QPOPOAVGASAATGPOPOQ--	ATQATGAPQG	252
63	Qy	PLAPPASPGP-----	FATRSPLETFMRPSSLLSRSSSGYFSTFDTDRSPAMSCDKSTQT	116	
253	Db	P--PGAAPGPPSPSPLRQPQNPQANPQLRSLLL-----	NPAP-----PQT	290	
117	Qy	PSPOQAFNHYL 1.28			
291	Db	GVPPOQASUHTL 3.02			

RESULT 15
 Q96VJ2 PRELIMINARY; PRT; 874 AA.
 ID Q96VJ2; PRELIMINARY; PRT; 874 AA.
 AC Q96VJ2; PRELIMINARY; PRT; 874 AA.
 DT 01-DEC-2001 (TREMBL; 19, Created)
 DT 01-DEC-2001 (TREMBL; 19, Last sequence update);
 DT 01-DEC-2001 (TREMBL; 19, Last annotation update)
 PROBABLE PROTEASE 1 LIKE PROTEIN.
 DE PROBABLE PROTEASE 1 LIKE PROTEIN.
 GN PCCM01B2..03.
 OS Pneumocystis carinii.
 OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
 OC Pneumocystis.
 OX NCBI_TaxID=4754;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PNEUMOCYSTIS CARINII F. SP. CARINII;
 RA Seeger K., Quail M., Harris D., Hall N., Wakefield A., Smulian A.G.,
 RA Submission M.T., Stringer J.R., Kieley S.P., Barrell B.G.;
 RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL592263; CAC42801.1; -.
 DR PROTEASE.
 KW PROTEASE.
 KW SEQUENCE.
 KW 874 AA; 97929 MW; Q8REF527C4D9603A CRC64;

Query	Match	9.3%	Score	99.5;	DB	3;	Length	874;
Best Local Similarity	27.7%	Pred. No.	0.17;					
Matches	36;	Conservative	8;	Mismatches	53;	Indels	33;	Gaps
Qy	5	PSDVYS-	-----	ECDRERQQLQPAERPQLRGAPTSLOTEPOGNPEGNHCCEG	52			
Db	677	PSDIDANPOPQDPGSPSSSDPSPSSDPSPSSDPGPSPS-	-----	-SDPEPSSEPGYQP	734			
Qy	53	DCSPHGSQGPL-	APPASPQGPATRSPLFIMRRLSSLLSRSRSSGYFSFDDTRSPAPMSC	110				
Db	735	PSEPP-PPQQPLDPRPKSKDPNPNSDP-----	-----	-----	777	SSQQDPDTLSUSSNPTST		
Qy	111	DKSTQTPSPPP	120					
Db	778	SSSEPPPPPP	787					

Search completed: August 16, 2002, 10:04:56
Job time: 234 sec